Characterizing Survey Response: Estimating Cognitive Proxy

Yeabin Moon *

June, 2022

Click here for the most recent version

Abstract

This paper examines the response behavior in the popular survey. There is a growing interest in the cognitive decline of the old population, but not enough is known about its consequences and implications. The major challenge is insufficient measures of one's cognition in most survey data. The objective of the paper is to estimate a cognitive proxy from a survey response. Note that taking a survey requires a series of cognitive tasks. I propose a standardized measure of characterizing responses to open-ended financial questions. The resulting proxy shows appealing characteristics and aligns with the cognitive measures directly available in the Health and Retirement Study. I apply the method to the Panel Study of Income Dynamics, and the proxy performs similarly.

JEL classifications: I10, C80, C83

Keywords: Cognitive declines, Survey response, Cognitive test measure

*I would like to thank Elaine M. Liu, Chinhui Juhn, Vikram Maheshri, and Aimee Chin for their insight and comments. I also thank the seminar participants of the micro workshop at the University of Houston for helpful comments and suggestions. All remaining errors are mine.

1 Introduction

Cognitive decline is defined as difficulty with the process of using the brain functions to consider something. It could occur gradually or suddenly, and it could be temporary or persistent. Cognitive decline is one of the most common conditions among the elderly. According to the public health issue report from the Centers for Disease Control and Prevention, every one in nine adults reports episodes of cognitive decline. The incidence prevalence is 10.8 percent among adults 45–64 years old and 11.7 percent among adults aged 65 years and older.[1] It is becoming a cause for public concern given that about 30 percent of adults experiencing cognitive decline live alone.[2] It is one of the earliest noticeable symptoms of Alzheimer's disease, and about 10–20 percent of people over 65 years of age with cognitive decline develop dementia (Jessen et al. 2014). A decline in cognitive ability might be expected as people age, but stressful life events or medical incidents could increase the risk of cognitive decline. For example, labor market experience would affect later cognitive function (Mayeda et al. 2020). In-utero exposure to maternal stress from various experiences would also affect a range of cognitive outcomes later in life (Persson and Rossin-Slater 2018; Barbosa-Silva, Santos, and Rangel 2018). A growing amount of evidence, albeit limited in its causal link, suggests further investigation of the cause and effect of cognitive decline, but the lack of data availability limits researchers' ability to investigate questions of high policy importance (Chandra, Coile, and Mommaerts, forthcoming).

In this paper, I propose a measure as a proxy of cognitive ability based on the response patterns of popular surveys. Participating in a survey requires mental effort. I assume that a person with low cognitive ability would round up the responses more or be more likely to opt out of questions compared to higher cognitive participants. Hence, the respondent experiencing a cognitive impairment would provide fewer significant digits of numerical answers or skip more questions than other respondents. I choose the open-ended financial questions in particular. The financial questions are available in popular survey-type microdata such as the Survey of Consumer Finances (SCF) and the Panel Study of Income Dynamics (PSID). Researchers use different household balance sheet data types to study individual saving dynamics or explain many other economic behaviors. Adding

1. See Subjective Cognitive Decline – A Public Health Issue (link.)

2. Seniors who live by themselves could be more susceptible to poor health outcomes than those living with others (Gibson and Richardson 2017; Portacolone et al. 2018).

a cognitive proxy to these data sets would help the researchers study more questions from a different perspective. Also, the literature finds that answering the financial questions demands advanced mental operations (Riddles et al. 2017). Answering open-ended questions requires understanding of question statements, the ability to search memory, and the competence to format the response on a given response scale. Financial questions tend to be more challenging, so financial literacy is often associated with cognitive abilities (Muñoz-Murillo, Álvarez-Franco, and Restrepo-Tobón 2020; Bucher-Koenen and Ziegelmeyer 2011; Cole and Shastry 2008). Hence, analyzing the response patterns of the financial questions would provide a better understanding of cognitive functioning.

For this project, I use two microdata sources: the Health and Retirement Study (HRS) and the PSID. Both data sets provide extensive longitudinal data on the U.S. population with a range of health and financial information. They have many open-ended financial questions, and the question formats do not dramatically change across the waves. The PSID sample represents the U.S. population more than the HRS sample, but the HRS has more observations of the old population. Also, the HRS has the questionnaires directly related to cognitive ability, so the performance of the proxy can be tested within the HRS data set. I estimate the proxy using the HRS data and apply it to the PSID data.

I select financial questions based on the response rate and analyze the responses. It is worth noting that the HRS provides opt-out options for respondents, such as *do not know*, and I find that many respondents choose opt-out options[3]. Also, I observe that many respondents give an approximate number, rounded to the precision of 1, for example, 3,000 or 10,000. This type of numerical response format is defined as the maximal rounding Gideon, Helppie-McFall, and Hsu (2017). Based on the observations, I assume that the response format is related to cognitive ability and that the difficulty of the mental load increases in the following order: choosing opt-outs, maximal rounding, and all other types of numerical response. I numerically characterize each response: 0 for opt-outs, 1 for maximal rounding, and 2 for numerical answers. I construct the proxy by taking an average of them.

I perform a range of tests to validate the proxy. Existing work in cognitive psychology shows that cognitive ability exhibits a decline with age (Kaufman and Horn 1996; Salthouse 2010). Hence, the proxy should be validated by its aging patterns. Given that the HRS has the questions directly

3. The opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it.

related to the respondents' memory and cognitive ability, the cognitive proxy could also be tested against these direct measures, at least for the older population. I am interested in understanding the association between the proxy and the demographic variables, but endogeneity is a potential concern. I employ the individual fixed effects to relax this issue. I apply a similar set of tests to the proxy constructed from the PSID samples.

Rounding or digit preference, also known as response heaping, is frequently observed in interview-administered surveys. Specifically, the responses to the financial questions are very approximate at best.[4] Uncertainty about the true response is a common problem when researchers analyze such a numerical question. There is widespread support for taking numerical responses at face value, but many researchers try to estimate the degree of heaping in data and develop techniques to handle resulting problems such as attenuation bias (Manski and Molinari 2010; Roberts and Brewer 2001; Zinn and Würbach 2016). I provide a way of measuring cognitive ability assuming that the response patterns are the consequence of the mental effort. Hence, employing the proxy would provide another way to handle such heaping patterns or accuracy problems in the survey data. The empirical investigation of the extent to which the survey response would predict the respondent's cognitive ability contributes to several distinct literatures. The first related body of work examines the survey completion of the entire survey (Krosnick 1991). A set of papers explore the proportion of the skipping pattern and study the reasoning behind the opt-out behaviors. Many surveys offer the respondents explicit opt-out options such as *don't know* or *refused to answer*. Colsher and Wallace (1989) show that the old respondents are more likely to choose opt-outs. Their findings explain that opt-out answers are considered as taking cognitive shortcuts to make question answering easier and that this pattern partly captures the cognitive decline. Knäuper et al. (1997) specifically study how often the respondents aged 70 years old and older choose *don't know* and find that the respondents with low cognitive ability tend to choose this option more on difficult questions. I use these literature findings to base my assumption on the relationship between cognitive ability and choosing the opt-out options. Another set of papers examine the relationship between cognitive ability and the pattern of numerical response. Holbrook et al. (2014) focus on the specific digit preference. Gideon, Helppie-McFall,

4. Riddles et al. (2017) examine the difference in financial response between the micro survey data and the IRS account.

and Hsu (2017) examine rounding behavior in a broader manner and study the intraclass variation of the respondent's rounding behavior. Andrews and Herzog (1986) examine the relationship between the variance of the numerical responses and the respondent's age. The findings indicate that the respondents more round their responses when facing more challenging questions and growing older.

To the best of my knowledge, there is no research on estimating the respondent's cognitive ability based on the response patterns. Based on the literature findings, one might raise the question of whether focusing on opt-out responses would be enough to capture the cognitive ability. Admittedly, the overall patterns of opt-outs could be a good approximation for cognitive ability, but the response rate for those who choose opt-outs is less than 20 percent in the HRS data. Given that most respondents answer in numerical formats at least one financial question in numerical formats, one might also wonder whether focusing on the level of rounding the numbers would be an alternative as a cognitive proxy. However, it is hard to assume that cognitive ability is a linear function of significant digits of the numerical answer. I construct two other proxies based on these claims and compare the performance against my proxy in the paper.

I find the following results from my empirical analysis of the proxy. First, the proxy starts to decline from the mid-60s. The most significant changes in cognition with aging would be a decline in performance on cognitive tasks (Murman 2015), but the effects should notably happen for older people. My proxy would not decline until the mid-60s and then progressively worsen with age. It happens to both HRS and PSID datasets. Second, the proxy is positively correlated with cognitive ability, directly measured by the HRS. As noted above, the HRS has some questions regarding cognitive ability, and the proxy captures it as well. My proxy has some advantages over the traditional proxies, such as the years of schooling or age. Although education level has predictive power on cognition, it does not capture the aging trend of cognitive ability because schooling years rarely change among the elderly. Using age as a proxy for cognition captures the linear age trend because age certainly changes over time. However, it implies that the marginal effect of age on cognitive ability is the same for everyone, which is too strong an assumption to be employed. My proxy not only presents the aging trend but also generates meaningful distribution among the respondents. I further examine the correlation between the proxy and other demographic variables to see whether it closely matches the literature findings. The proxy captures both the aging pattern and the cognitive measurement directly from the HRS, even after controlling for the individual fixed effects.

The remainder of the paper is organized as follows: Section 2 presents the data description for the HRS. Section 3 lays out the way of the proxy estimation. This section describes how to characterize the response and evaluate the resulting proxy. Eventually, I want to apply the method to other survey data that do not have any cognition measurements and therefore give a guideline for making a cognitive proxy for the researchers who want to study cognition but do not have it in their data set. In section 4, I test whether Title VI of the 1964 Civil Rights Act would impact the cognitive formation of African Americans in the Southern U.S using the PSID as an application of this project. Section 5 concludes.

2 Data Description

The Health and Retirement Study is a multi-dimensional household longitudinal data surveyed by the Institute for Social Research at the University of Michigan. It enables researchers to analyze U.S. seniors' individual (or household) behaviors.I focus on the 2004 to 2018 waves of the Public Survey Data with the intention of adding the cognitive proxy to this data. The sample is limited to the household head aged between 50 and 89 during the survey periods. The HRS contains a set of demographic variables such as gender or educational achievement, a range of socioeconomic indicators, and financial portfolios. Note that the sample in the HRS is built up over time. I start with the 2004 survey because it is the first year that the baby boomer generation is included. Some financial questions are asked at the household level, and the HRS classifies the individual respondent who responds to the question on behalf of the household.[5] In most cases, one family member answers every household level question. If the roles are separated among the family members, I mainly focus on the individual responding to the financial questions because the interest lies in the open-ended financial questions.[6]

The sample contains 62,851 individual-years, pooling 16,187 individuals across the seven survey

5. The HRS classifies every individual respondent as the financial respondent, family respondent, or cover screen respondent (the first respondent interviewed).

6. The observations are also dropped if the primary respondent is either sibling, child, or helper.

waves. Table 1 presents the parsimonious summary statistics of the data. The overall demographic does not change by the survey year or age group. The respondents are mostly female and achieve high school degrees on average. Many respondents encounter different modes of interviews during the periods. Before 2002, the HRS is surveyed through face-to-face interviews at baseline, and the telephone is used for follow-up interviews. The selection issue in the interview mode is not a concern because the interview mode is randomly determined after 2006. One possible concern is the lack of racial information. The HRS oversamples African-American and Hispanic households at about twice the rate of Whites, but race information is restricted to public data (Sonnega et al. 2014). It will raise a concern if the race is a crucial determinant of the outcome of interest.

2.1 Question selection criteria

The HRS contains a range of financial questions.[7] I select the questions based on the following criteria. First, the response format should be open-ended and require more than 2-digit numbers. It rules out multiple-choice answers, so the set of answers is either the numerical value or the opt-out responses such as *don't know* and *refused to answer*. For the questions with less than a 2-digit number, they are also dropped because it is hard to decide whether a respondent approximates it or not.[8] Second, the response rate should be significant. This is mainly because some questions are only applicable to a certain group of respondents–such as the amount of alimony–so their representativeness is in question. Also, the low response rate would mean that the question is of little value to consider. Third, the selected questions should appear in multiple waves to analyze the patterns over time. During the period of research interest, the ten questions meet the criteria: *the present value of home, real estate taxes, social security income, checking account, transportation value, food at home, food away from home, out-of-pocket for doctor visit, dental bills* and *drug costs*.[9]

One common characteristic among the ten questions is the digit preference for round numbers.[10]

7. Financial questions mainly belong to the following sections: health care cost (N), family structure (E), housing (H), asset and income (Q), capital gains (R), job (J, L), disability (M), insurance (N), and divorce (S).

8. Time or date-related questions are also out of the list.

9. Note that these questions are self-reported information. The HRS links the public data to the Social security earning data and the Medicare records, so the actual information on some questions is available in the restricted data. This project utilizes publicly available data sources only.

10. See Figure 1.

The tendency is particularly strong on the maximal rounding. Note again that the maximal rounding indicates the numerical response format in which the number is rounded to the precision of 1. That is, the leftmost digit is any number, and the rest of the digit positions are zeros, for example, 2,000 or 100,000. Table 2 presents the summary statistics for the ten questions regarding maximal rounding and opt-out responses. It indicates that more than one-third of variations come from either the maximal rounding or the opt-outs responses. I leave further analysis of these questions in the later section.

2.2 Cognitive score

The HRS has a series of tests of one's cognitive ability. The questions are related to episodic memory, mental status, and comprehension of vocabulary. Each correct answer scores numerical points. I define the *cognitive score* as the sum of the three questions, *immediate word recall*, *delayed word recall*, and *serial 7's test*, and use it as a reference index of the cognitive ability.[11] For the immediate word recall, the respondents observe the ten-noun list and then are asked to recall the list in order. A few minutes later, the respondents are asked to recall the same list of the words again. This is the delayed word recall. For the serial 7's test, the respondents are asked to subtract 7 from 100, and continue subtracting 7 from the last response. The respondents are asked to try this five times. The score range of the immediate word recall and delayed word recall are both zero and ten, and the serial 7's test is between zero and five. Therefore, the cognitive score ranges from zero to twenty-five.

Figure 2 plots the average score of the cognitive questions against age in years. All the panels show a similar aging trend. There is little variation until age 65, but sharp declines are observed after 65. There is no consensus on when the cognitive decline starts in medical literature, but there is a general agreement that the cognitive decline would be accelerated for the old people (Karr et al. 2018). Hence, the cognitive score will serve as a benchmark for the proxy.

11. Herzog and Wallace (1997) claim that the selected questions are relevant to verbal learning, reasoning, and attention abilities, which are essential to cognitive skills. Weir (2017) surveys the entire list of cognitive measures.

2.3 Demographic variables

Demographic variables include age, mode of interview, gender, years of schooling, and wealth. In the later section, I test whether the correlations between the proxy and these controls are consistent with the literature. Previous findings indicate that family wealth would affect cognitive development in most stages of life (Schady et al. 2015; Cagney and Lauderdale 2002). I measure the family wealth using the sum of nonfinancial wealth, retirement wealth, and other financial wealth. Since each value is nominal, I first deflate the values using the price index in 2012 for Gross Domestic Product from the National Income and Product Accounts. Race and geolocation would be important factors, but I do not use the masked variables. In many data, these variables are considered individual identifiers, so most users would not have access. Note that the goal of this project is to find an accessible way to measure cognitive ability. But if available, they should be taken into consideration.

3 Proxy Estimation

The respondents in the HRS either provide numerical values or opt out of the open-ended questions. Providing incomplete answers such as round numbers or skipping the questions are called *satisficing* (Krosnick 1991). Because the difficult questions would cause respondents to avoid cognitive efforts, the satisficing response could be the result of individual optimization. Krosnick (1991) claims that the satisficing behavior is negatively correlated with cognitive ability controlling for the difficulty and the motivation to participate in the task. The following subsections investigate whether the cognitive ability could be inferred from the response patterns. I propose a standardized method that enables cross-question comparisons in answering behavior.

3.1 Characterizing the response

The numerical response to the open-ended financial questions approximates the actual value at best (Gideon, Helppie-McFall, and Hsu 2017). Assuming that responding to financial questions with round numbers is a consequence of satisficing, the level of rounding could be used as a proxy for cognitive ability. Gideon, Helppie-McFall, and Hsu (2017) define the level of rounding in the following way. For a given numerical value, they count the number of significant digits, n, and the number of total digits, m.[12] The level of round is measured by $\frac{m-n}{m-1}$. For example, the level of round of 30,000 and 33,300 are 1 and 0.5, respectively. This index ranges between zero and one. The higher the level of round, the more the respondents round up the responses. Figure 3 presents the average level of round against age for every question.

Note that it is not clear whether the level of round has an increasing aging pattern in Figure 3. Table 2 indicates that a significant number of respondents present the maximal rounding responses in every question. It raises the question of whether it would be sufficient to focus just on the maximal rounding instead of using the level of round to capture cognitive ability. Figure 4 presents the maximal rounding on average against age for every question.

Both the level of round and the maximal rounding would help characterize the numerical response. However, the HRS provides the opt-out options such as *don't know* and *refuse to answer*. Note that the literature indicates that the reductions in cognitive functioning can be reflected in the completion rate of the survey, and the opt-out response is one of the main examples (Knäuper et al. 1997). Using a dummy variable that captures the opt-out response, Figure 5 presents the opt-out responses on average against age for every question.

Ideally, the proxy should be responsive to the level of round on numerical response. At the same time, it should capture whether the respondent chooses the opt-out options. I suggest a way to construct the cognitive proxy to consider both aspects. I divide the response format into three, opt-out, maximal rounding, and numerical answer, and assign zero, one, and two, respectively. I assume that choosing opt-outs demands the least cognitive functioning ability and that any type of numerical answer other than maximal rounding is the most demanding work. For comparison, I employ two other ways of classifications. The first classification is the level of round defined by Gideon, Helppie-McFall, and Hsu (2017). It implicitly assumes the level of round is correlated with cognitive ability. The second classification is the binary variable indicating whether one chooses opt-out responses. To the best of my knowledge, these two methods are never evaluated as a proxy for

12. The significant digits (or significant figures) for the natural numbers are defined as all non-zero numbers or zeros between non-zero numbers. The trailing zeros are not considered as significant.

cognitive ability.[13] I simply borrow the method of characterizing the responses. After characterizing the response using these three ways, I take the average of them. For each proxy, I label them as *Moon, Gideon* and *Knäuper*. I exclude social security income and vehicle value questions from the list. Social security income is mostly applicable to respondents over 65, and the vehicle question is rather subjective and not timely. I will test them in the robustness check.

Importantly, I want the higher value of the proxies to imply higher cognitive ability. *Gideon* is the average of the level of round, so the higher the value implies a lower cognitive ability. Therefore, I use $1 - \frac{m-n}{m-1}$ when characterizing the response. Similarly, for *Knäuper*, I assign 0 to opt-out responses and 1 to numerical responses when characterizing each response. I standardize all three proxies.

3.2 Proxy Evaluation

In this section, I evaluate the proxy based on the two aspects. First, the proxy needs to have an aging pattern. The cognitive decline usually occurs with normal aging, but there is no consensus on when it accelerates. Figure 2 shows a kink in each cognitive measure directly available from the HRS questions. Each measure presents a sharp decline after the mid-60s. The ideal proxy should have a similar aging trend. Secondly, the proxy should positively correlate with the cognitive measures. I perform the regression analysis in this regard with demographic controls.

3.2.1 Trends in Aging

Figure 6 plots each proxy against age in years. I benchmark its performance against the cognitive score in the first panel. The proxy *Moon* seems to capture the aging pattern reflected in the cognitive score closely. The mean value would not change until the mid-60s, after which it would fall. Whereas the mean value of proxy *Knäuper* progressively decreases with age. Given that the proxy *Knäuper* characterizes the response solely on the opt-out options, it is appealing for the sake of simplicity. However, its capability is in question for the younger cohort. The proxy *Gideon* looks counter-intuitive against the assumption that the respondents would more likely round the numerical answers

13. Gideon, Helppie-McFall, and Hsu (2017) investigate the potential cause of the digit preference, and Knäuper et al. (1997) examine the cause of opt-out responses.

as they age. However, this method ignores opt-out responses. Since the proxy *Gideon* is just an average of the level of round, the increasing trend would not necessarily imply a more rounding trend. Because Figure 6 shows the cross-sectional average by age group, the patterns are susceptible to endogeneity issues. To relax this concern, I evaluate the aging pattern of the proxies controlling for individual (non-time varying) effects. I partial out non-time varying controls within the same respondents and examine the residuals against age. Figure A.1 presents the corresponding residual plots. Note that *Gideon* would not present increasing patterns anymore. Interestingly, *Moon* and *Knäuper* now present similar patterns. The proxy *Moon* assumes that the maximal rounding implies a less cognitive activity from the other type of numerical response. Then, Figure A.1 would imply that the resulting aging patterns are not sensitive to the assumption if longitudinal data is available. However, given that more than 60 percent of the respondents provide numerical answers, employing more variations from the numerical responses would provide a more flexible fit when constructing a cognitive proxy in a cross-sectional setting.

3.2.2 Regression Analysis

This subsection investigates the association between the proxies and a range of demographic variables in detail. The proxies serve as the outcome of interest. The cognitive score and age serve as the main independent variables. Note that cognitive development would be affected by income, education, and gender (Myers 1976; Budd and Guinnane 1991; Boyle and Gráda 1986). I perform the regression analysis on the proxies to further evaluate whether the resulting associations are consistent with the literature.

Consider the following regression: For the respondent i in the survey year t,

$$Proxy^{j}_{i,t} = \alpha^{j} + \beta^{j} cognitive\ score_{i,t} + f\left(age_{i,t}\right) + X_{i,t}\gamma^{j} + \delta_t + \epsilon^{j}_{i,t} \tag{1}$$

where $j = \{Moon, Gideon, Knuper\}$ indicates the proxy classification. The regression model tests the partial correlation between the proxies and the selected demographic variables available in the HRS. $Proxy^{j}_{i,t}$ is the value of the corresponding proxy for i at t. $cognitive\ score_{i,t}$ is the sum of the scores from the cognitive questions available in the HRS for i in the survey year t. $age_{i,t}$ indicates i's age at t based on the birth date. Note that the answering behavior, or cognitive ability in general, is a function of age, and hence, the relationship with age is an important parameter to measure the performance of the proxies. $X_{i,t}$ contains the set of the demographic controls for i at t, including the binary indicators for the phone-interview mode, high school graduation, gender, and wealth quantiles. δ_t captures the survey-year fixed effects.

Table 3 presents the estimation results. The column labels denote the proxy classification. All regressions restrict the HRS household heads aged between 50 and 89 who participated in the survey from 2004 to 2018. I begin with a relatively parsimonious specification assuming age and the cognitive proxy have a linear relationship in column 1. An increase in one standard deviation of the cognitive score is associated with an increase in about 0.09 standard deviation of the proxy value. The additional year effect of age would lead to a decrease in about -0.005 standard deviation of the proxy. Both estimates are statistically significant at any conventional level. A key finding is that even in the parsimonious specification, *Moon* is positively correlated with the cognitive score and negatively correlated with age, as expected.

Motivated by Figure 6, I estimate the following specification:

$$\begin{aligned}\text{Proxy}^j_{i,t} &= \alpha^j + \beta^j_1 \text{cognitive score}_{i,t} \\ &+ \beta^j_2 \mathbb{1}\left(60 \le \text{age}_{i,t} < 70\right) + \beta^j_3 \mathbb{1}\left(70 \le \text{age}_{i,t} < 80\right) + \beta^j_3 \mathbb{1}\left(80 \le \text{age}_{i,t}\right) \\ &+ X_{i,t}\gamma^j + \epsilon^j_{i,t}\end{aligned} \tag{2}$$

Instead of adding the continuous age variable, it includes three age dummies spanning ten years each. The respondents aged between 50 and 59 serve as the base level. Column 2 uses *Moon* proxy, and the correlation with the cognitive score would not change much against column 1. The effect of the first age dummy indicates that between 60 and 70 is not statistically significant, but the two other dummies get stronger negatively and statistically. The result corresponds to the aging trends in Figure 6. The estimates in column 3 show again the poor performance of *Gideon* in the cross-sectional settings. The estimates in column 4 show a similar pattern compared to column 2. I estimate the regression again with more age dummies in Table A.1. The marginal effect on the age begins much earlier on average when using the proxy *Knäuper* compared to *Moon* as in Figure 6.

Of course, the resulting estimates are susceptible to the omitted variable bias.[14] To address this concern, I perform the regressions again, adding the individual fixed effects in Table A.2. Note that the regressions use the continuous age variable here. Interestingly, I find that the coefficients on the cognitive score are statistically positive in columns 1 and 3 with similar magnitude, but that the marginal effect of age is different. Proxy *Moon* shows that it remains statistically negative, but there is no statistical effect on proxy *Knäuper*. Based on the findings, characterizing the maximal rounding into the proxy construction would be the preferred choice because it performs reasonably well both in cross-sectional and longitudinal settings.

3.3 Robustness check

One worry is that the question-selection criteria for the proxy are rather subjective. Hence, I need to test whether the findings above would be sensitive to the selection scheme. Recall that seven questions are used. One would raise the question of whether one or two responses would have more significant influences than others, and hence it would result in a systematic bias on the proxy. To check this claim, I subtract one question from the seven questions and evaluate the resulting *Moon* proxy. Table 4 presents the estimation results using the same specification (2), and the column label indicates what question is left. Any estimates within Table 4 would not stand out, and there is no significant difference compared to column (2) in Table 3.[15]

Another concern would be a shortage of questions for the proxies. There are only seven questions selected, so the proxies would not capture the important variations from the questions not selected. To alleviate this concern, I add two more questions: the social security income and the value of vehicles. The response format of these questions is open-ended, and the response rates are above the criteria. However, the social security income is only applicable to a certain population, and the vehicle question is not directly related to memory.[16] I use the same specification (2), but the sample

14. In fact, race and geolocation are essential attributes of cognitive development (Rushton and Jensen 2005).
15. I try out the same evaluations on the other proxies but find no meaningful difference.
16. Also, the vehicle-value responses have inconsistencies within the same person over time, especially for the multiple

is further limited to the respondents aged above 65 years old. Table 5 presents the results. Columns 1 and 3 use the proxy *Moon* with two more questions. Columns 2 and 4 use the original set of questions as a comparison. Columns 3 and 4 further include the individual fixed effects. Given that there is no meaningful difference between column 3 and column 4, the question selection schemes would be insensitive to the longitudinal data. However, adding the two questions would result in negative estimates on the effect of high school graduation in column 1, which is counter-intuitive. Because of endogeneity, admittedly, it is hard to conclude which set of questions are preferred here, but it is advisable to use the original questions, at best.

4 Validating the Proxy: Application to the PSID

The goal of the project is to construct a proxy from the way people respond to survey questions even when a direct measure of cognition is not available. I use the HRS to construct and evaluate the proxy based on the cognitive score directly available in the HRS questionnaires. However, without high external validity, it is hard to generalize the findings of this project. Hence, in this section, I use the Panel Study of Income Dynamics (or PSID) to extend the scope. There are several reasons why I choose the PSID. First, it has similar features compared to the HRS. The PSID has a panel dimension, a similar set of questions, and provides opt-out options for each question, such as *don't know* and *refused to answer*. So, it is possible to examine whether the proxy presents similar patterns compared to the HRS. Second, the PSID collects data far earlier than the HRS so that the proxy can be tested against the stylized historical findings. The PSID is especially useful since it has more demographic variables such as race and birth location.

4.1 Proxy estimation and evaluation

The PSID sample consists of three respondent groups: the core sample, low-income-family sample, and immigrant sample. The core sample is the largest, representing the US population, so I focus on

car owners.

the household head from the core sample to match the HRS sample.[17] Similar to the HRS data I previously used, the sample is limited to the respondents aged between 50 and 89 who participated in the surveys from the 1999 to 2019 waves. It represents 32,828 household head-years, and 6,773 unique respondents participate in the surveys. One concern is that the question-selection criteria from the HRS are hardly applicable to the PSID because missing or unreliable responses are already imputed in some questions. Hence, I select the following questions: *home value, labor wage, monthly rent, property tax, food home,* and *food out*. I try to select a similar set of questions in the PSID. These questions are not imputed, and the response rates are high. The full description of the questions and the summary statistics are presented in Table 6.

I construct the three proxies like before and standardize them. The PSID does not have direct measures or questions regarding cognitive ability, so there is no benchmark to compare their performance within the PSID.[18] I evaluate the proxies based on the two aspects similar to section 3.2. Figure 7 presents the cross-sectional age trends of each proxy. *Moon* and *Knäuper* seem to capture the aging patterns presented in the cognitive score of the HRS. Again, *Gideon* shows the counter-intuitive pattern, and hence, the level of round would not be a good proxy in cross-sectional settings. Figure A.5 presents the residual plots after controlling for the individual fixed effects. Similar to the findings in section 3.2, *Moon* and *Knäuper* present decreasing patterns. Admittedly, the aging trend in *Knäuper* looks more appealing. However, given that more than 70 percent of respondents provide numerical answers in the PSID, its applicability is in question.

I further investigate the association between the proxies and a set of variables available in the PSID. Although the PSID does not have cognitive measures, it provides race and home location information at the state level, which are essential attributes of cognitive development (Rushton and Jensen 2005). I perform the regression analysis similar to the specifications (1) and (2). Table 7 presents the estimation results. The proxy *Moon* serves as the outcome of interest. All the regressions include the indicators of high school graduation, female, wealth quantile, and survey-year fixed effects. Columns 1 and 2 use a similar set of controls as in the HRS analysis, except for the cognitive

17. However, it does not necessarily mean the two samples are comparable. The HRS oversamples African-American and Hispanic households at about twice the rate of Whites, but the race information on the respondents is masked.

18. One exception is the sentence completion quiz in the 1972 wave, but it is never repeated.

score and the interview mode.[19] Column 3 further includes race and home location information. All the estimates would not stand out, and the other two proxies also present similar results.[20] Still, the marginal effect of the age would get stronger with age, and hence the non-linear specification would have a better fit against the linear-age model. The effect of being black is negatively related to the proxies and is statistically significant at any conventional level. It is intuitive since black people would have less access to the resources that aid cognitive development, such as education and nutrition, and be more susceptible to adverse environments. Table A.4 presents the estimation results with the individual fixed effects. Interestingly, I still find that the marginal effect of age on proxy *Moon* remains statistically negative, but there is no statistical effect on *Knäuper*. This is a similar finding in Table A.2, and hence most findings in the HRS are applicable to the PSID.

4.2 Impacts of 1964 Civil Rights Act

In this section, I evaluate whether Title VI of the 1964 Civil Rights Act would have an impact on the cognitive formation of African Americans in the Southern U.S. Black people have had the worst health outcomes of all major demographic groups (Nelson 2002). Many features contribute to these disparities, and the limited access to medical care is one of the leading sources (Cutler, Lleras-Muney, and Vogl 2008). Given that there are many channels explaining the access gaps, such as socioeconomic status or mistrust, Almond, Chay, and Greenstone (2006) specifically examine the segregation of medical access in the Southern states. They observe that the significant racial segregation across hospitals existed in southern rural states in the mid-1900s and find that the federally mandated integration of health care facilities improved health outcomes for the black people. They analyze the trends in black-white infant mortality gaps in Southern states between 1955 and 1975, and find that the gaps were closing in the rural south, where the segregation was most pervasive before the policy.

I revisit their findings and investigate whether the de-segregation policy would improve the cognitive ability of the corresponding African Americans based on the fetal origins hypothesis. The hypothesis, also known as the Baker Hypothesis, claims that the fetal conditions would permanently

19. The interview mode information is available in the PSID, but more than 95 % of respondents participate in the survey through the phone.

20. See Table A.3.

affect later-life health outcomes, and a growing literature finds that it would affect the socioeconomic status as well as labor-market outcomes in a later stage of life (Almond and Currie 2011; De Boo and Harding 2006). If the federal mandate caused the black-white infant mortality gap to fall, it would likely mean that the access to prenatal care would increase for the black mothers in the South. The fetal origins hypothesis predicts that the improvement in access would result in a different trajectory of cognitive development for the black residents born after the policy in the South.

I use the 1999 to 2019 waves of the PSID same as before, but the sample is further limited to those born between 1955 and 1974 in the Southern states. Table 8 summarizes the descriptive statistics. The first two columns present the average demographic variables for this sample, disaggregated by the enactment of the integration policy. In general, African-American respondents are over-sampled in the southern regions.[21] I further disaggregate the samples based on whether the grown-up area is an urban area or not in columns 3–6 and find no significant change in the demographics after the integration policy. One concern is that the performance of the proxy is tested only on the old population. It is because the benchmark cognitive score from the HRS is applicable to the respondents aged over 50. Figure A.6 shows the average value of *Moon* for the respondents between 20 and 89 participating in the surveys from 1999 to 2019, and it looks to capture well the life-cycle patterns of cognitive development.

I test the following simple trend break regression model based on Almond, Chay, and Greenstone (2006). For the respondent i in the survey year t,

$$\begin{aligned} y^j_{itsu} &= \beta^j_1 \mathbb{1}\left(50 \leq age_{i,t} < 60\right) + \beta^j_2 \mathbb{1}\left(60 \leq age_{i,t}\right) + \beta^j_3 \cdot age_{i,t} \cdot \mathbb{1}\left(i \text{ born} > 1965\right) \\ &+ X_{it}\gamma^j + \alpha^j_t + \alpha^j_s + \alpha^j_u + \epsilon^j_{itsu} \quad \text{for } j = \{\text{rural black, rural white, urban black, urban white}\} \end{aligned} \tag{3}$$

The proxy *Moon* serves as the outcome of interest based on the responses to the PSID questions. Let s and u further denote the grown-up state and the urban/rural index. Motivated by the previous finding, two age dummies are added along with the trend break interaction term, indicating whether i born after the integration policy would have different cognitive development over age. X_{it} includes

21. The black residents especially dominate the rural area. For example, the share of African American respondents is more than 80 percent in most Alabama and Mississippi states.

the indicators of female, high school graduation, and wealth quantile. I employ the survey year fixed effects, the grown-up state fixed effects, and the grown-up rural indicator fixed effects. Standard errors are clustered at an individual level. The sample is disaggregated into four groups, *rural black, rural white, urban black*, and *urban white*, and equation 3 is estimated separately by this group. According to the hypothesis, $\beta_3^{\text{rural black}}$ would stand out, indicating that the black people in the rural area of southern states born after the policy would have better cognitive development than the black people in the rural area of southern states born before as they grow.

Table 9 presents the estimation results. Columns 1 and 2 use the black respondents who grew up in rural and urban areas, respectively. Columns 3 and 4 use white respondents who grew up in rural and urban areas, respectively. The coefficient on *post-1965 trend break* is of interest, and it is statistically positive only in column 1. The marginal effect of age would lead to an increase in *Moon* by about 0.03 standard deviation for black people in the rural south born after the policy. Note that the urban black samples would not have a significant gain from the integration policy. It affirms the hypothesis because segregation in an urban area would not be prevalent compared to rural areas even before the policy.

To evaluate the findings above, I focus on the three important variations. The first variation is from the integration policy. After 1964, the integration policy is adopted at the federal level. Since it targets to reduce the segregation in medical access to African Americans, race is the second variation. Note that the level of segregation differs by location and that it is particularly severe in the rural areas in the South. So, the location is the third variation. Motivated by Muralidharan and Prakash (2017), I employ the triple-differences estimation to measure the effects of the integration policy in the South as a robustness check. Consider the following specification. For the respondent i in the survey year t who grow up in the state s,

$$\begin{aligned} y_{its} &= \gamma_1 post_i + \gamma_2 rural_i + \gamma_3 black_i \\ &\quad + \gamma_4 \left(post_i \cdot rural_i\right) + \gamma_5 \left(post_i \cdot black_i\right) + \gamma_6 \left(black_i \cdot rural_i\right) \\ &\quad + \beta \left(post_i \cdot rural_i \cdot black_i\right) + X_{it}\delta + \alpha_t + \alpha_s + \epsilon_{its} \end{aligned} \tag{4}$$

The proxy *Moon* serves as an outcome. Note that post_i corresponds to $\mathbb{1}\left(i \text{ born} > 1965\right)$. $rural_i$

and $black_i$ indicate whether i grew up rural area and black, respectively. X_{it} includes the indicators of female, high school graduation, age dummies, and wealth quantile. The specification further includes the survey-year and grown-up state fixed effects, and the standard errors are clustered within an individual level.

Table 10 presents the estimation results. The coefficients on the controls are suppressed. The proxy *Moon* serves as an outcome in column 1. Similar to the previous findings in Table 9, the coefficient on the triple differences is statistically positive. The average effect of being black born after the integration policy and growing up in the rural south area raises the proxy by about 0.2 standard deviations, with no other groups increasing after the policy. Columns 2 and 3 use whether someone has diabetes or chronic ear problems before the age of 17 to see if the integration policy affects other health outcomes. The hypothesis claims that if the mother had better prenatal care, the child would have better health outcomes. In these two categories, the black respondents growing up in the rural southern regions after the policy would have had better conditions.[22] One might be concerned that the increase in the cognitive proxy may result from a higher educational achievement from Table 8. Column 4 uses the binary variable indicating high school graduation as an outcome instead of using it as a control. Contrary to the previous findings, no groups would benefit from the policy in terms of educational achievement. Note that the policy intends to target integrating racial mix, not promoting schooling years. It is true that the years of schooling increase in general, but the estimated result would imply that after the policy, black people would not graduate from high school at a higher rate than other groups. Therefore, this finding implies that the increase in cognitive proxy among the black people in the rural south during this period would not be from the increase in educational achievement. The more compelling approach would need county-level data. Note that the compliance rate varied among the hospitals and even within hospitals (Brown et al. 1999). Almond, Chay, and Greenstone (2006) use Medicare certification dates for Mississippi hospitals and analyze the impacts of the presence of Medicare-certified hospitals in a county on the health outcomes of the surrounding regions. Unfortunately, the county information is masked, and the least unit of the geocode is the urban/rural indicator in the public PSID data.

22. In fact, I test 19 categories of health outcomes. There are no statistically significant estimates on the triple difference term other than diabetes and chronic ear problems.

5 Conclusion

In this paper, I proposed the method of estimating cognitive ability in popular survey data. I presented how to characterize the open-ended financial questions into three forms: opt-out response, maximal rounding, and numerical response, assuming that they represented the respondent's cognitive ability in that order. The proxy based on the HRS data captured the aging trends, and it was positively correlated with the cognitive measurement directly available in the HRS questions. I applied the same method to the PSID data set, and it performed similarly. I also tested two other proxies and found that my method performs better in cross-sectional and longitudinal settings. As an application, I examined the impact of the 1964 Civil Right Act on the cognitive development of African Americans in the South and found that the results aligned with Almond, Chay, and Greenstone (2006). Based on its validity, I hope that researchers will use their data to test more questions about cognitive development.

Table 1: Descriptive statistics for the HRS respondents

variable	overall	<65	≥ 65
age	68.5	57.9	75.3
female	0.73	0.69	0.75
year of education	12.42	12.87	12.13
phone interview	0.45	0.46	0.44
total wealth ($1,000)	81	71	87
observations	62,851	24,801	38,050

The table summarizes the pooled sample of the household heads who participated in the HRS survey from 2004 to 2018. The sample is limited to the respondents aged between 50 and 89. Column 1 includes all the corresponding respondents. Columns 2 and 3 disaggregate this sample by age. Phone interview indicates the proportion of respondents who are offered phone interviews. The remaining respondents have a face-to-face interview. Total wealth is the sum of nonfinancial wealth, retirement wealth, and other financial wealth. The total wealth denotes the average of 1000s of 2012 dollars.

Table 2: Descriptive statistics for the selected questions in the HRS

variable	maximal rounding (%)		Opt-out (%)	
		Don't know	Refused to Answer	skip
Home value	30.9	15.6	1.2	11.9
Property tax	26	18.8	1.4	1.2
SSI income	22.6	12.5	.8	8.3
Checking	48.2	11	10.5	.8
Vehicle	53.1	17.7	1.4	.8
Food home	61	9.9	1	.6
Food out	42	3.1	.7	.5
OOP Doc	51.9	18.1	.6	1.2
OOP Dent	59.9	7.5	.5	1.1
OOP Drug	45.3	11.8	.4	5.6

Each row summarizes the response formats. The maximal rounding indicates the numerical response format in which the number is rounded to the precision of 1. That is, the leftmost digit is any number, and the rest of the digit positions are zeros, for example, 2,000 or 100,000. The respondents in the HRS can opt out of the questions, choosing *Don't know, Refused to answer*, or skipping the question. The column *skip* counts the missing responses only for those who have access to the question but skip it. The access is determined by the cross-references Core interview content (link). The sample is limited to the household heads aged between 50 and 89 who participated in the survey from 2004 to 2018. The Social security income is available to those above 60, so the third row further restricts the sample aged between 60 and 89.

The full question texts are:
Home value: *What is its present value? I mean, what would it bring if it were sold today?*
Property tax: *What were the real estate taxes in (LAST CALENDAR YR CALCULATED) on this home?*
SSI income: *About the Social Security income that you (yourself) receive, how much was that Social Security check, or the amount deposited directly into an account, last month?*
Checking: *If you added up all such accounts, about how much would they amount to right now?*
Vehicle: *What are they worth altogether, minus anything you still owe on them?*
Food home: *How much do you (and other family members living there) spend on food that you use at home in an average week?*
Food out: *bout how much do you spend eating out in a typical week, not counting meals at work or at school?*
OOP Doc: *About how much did you pay out-of-pocket for doctor or clinic visits?*
OOP Dent: *About how much did you pay out-of-pocket for dental bills?*
OOP Drug: *On average, about how much have you paid out-of-pocket per month for these prescriptions?*

Table 3: Regression analysis on the proxies

	(1) Moon	(2) Moon	(3) Gideon	(4) Knäuper
cognitive score	0.0933***	0.0919***	0.0109**	0.0756***
	(0.00448)	(0.00449)	(0.00490)	(0.00396)
age	-0.00589***			
	(0.000398)			
$60 \leq$ age < 70		0.00517	0.0723***	-0.0497***
		(0.0103)	(0.0112)	(0.00908)
$70 \leq$ age < 80		-0.0504***	0.114***	-0.157***
		(0.0107)	(0.0116)	(0.00941)
$80 \leq$ age		-0.175***	0.124***	-0.322***
		(0.0126)	(0.0138)	(0.0112)
phone	-0.00811	-0.0108	-0.0191**	-0.000973
	(0.00783)	(0.00783)	(0.00855)	(0.00691)
high school	0.00611	0.0113	-0.0747***	0.0317***
	(0.0106)	(0.0106)	(0.0116)	(0.00937)
female	-0.0168**	-0.0200**	0.0486***	-0.0459***
	(0.00842)	(0.00841)	(0.00919)	(0.00742)
wealth Q2	0.0450***	0.0451***	-0.123***	0.0440***
	(0.0121)	(0.0121)	(0.0132)	(0.0106)
wealth Q3	0.139***	0.138***	-0.178***	0.100***
	(0.0123)	(0.0123)	(0.0134)	(0.0108)
wealth Q4	0.233***	0.231***	-0.199***	0.162***
	(0.0125)	(0.0125)	(0.0136)	(0.0110)
wealth Q5	0.313***	0.308***	-0.257***	0.241***
	(0.0128)	(0.0128)	(0.0140)	(0.0113)
mean (proxy)	.1742	.1742	-.0158	.1920
N	46162	46162	45841	46162
adj. R^2	0.046	0.047	0.016	0.062
F	244.7	203.7	61.98	260.1

All regressions restrict the HRS respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. All the regressions include cognitive score, indicators for phone-interview mode, high school graduation, female, and wealth quantile as a set of controls. Also, all regressions include survey-year fixed effects. The dependent variables are the cognitive proxies defined below. All the proxies are standardized. Column 1 controls for age as a continuous variable. Motivated by Figure 6, columns 2–4 include three age dummies instead of adding the age variable. Each dummy spans ten years, and the respondents aged between 50 and 89 serve as the base level. Standard errors are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it. For more details, see Table 2.

Table 4: Robustness Check I: one question is subtracted from the proxy

	(1) home value	(2) pro tax	(3) checking	(4) food at home	(5) doctor	(6) dental	(7) drug
cognitive score	0.0826***	0.0756***	0.0950***	0.0920***	0.0897***	0.0897***	0.0878***
	(0.00457)	(0.00451)	(0.00474)	(0.00468)	(0.00450)	(0.00450)	(0.00452)
$60 \leq$ age < 70	0.00898	-0.0179*	0.00369	0.0139	0.0181*	0.0181*	-0.000369
	(0.0105)	(0.0103)	(0.0109)	(0.0108)	(0.0103)	(0.0103)	(0.0104)
$70 \leq$ age < 80	-0.0344***	-0.0887***	-0.0565***	-0.0342***	-0.0293***	-0.0293***	-0.0496***
	(0.0109)	(0.0107)	(0.0113)	(0.0111)	(0.0107)	(0.0107)	(0.0107)
$80 \leq$ age	-0.144***	-0.207***	-0.191***	-0.152***	-0.145***	-0.145***	-0.165***
	(0.0129)	(0.0127)	(0.0134)	(0.0132)	(0.0127)	(0.0127)	(0.0127)
phone	-0.0126	-0.0153*	-0.00842	-0.00685	-0.00868	-0.00868	-0.0107
	(0.00798)	(0.00788)	(0.00827)	(0.00816)	(0.00785)	(0.00785)	(0.00788)
high school	-0.0150	0.0247**	0.0231**	-0.00192	0.0125	0.0125	0.0137
	(0.0108)	(0.0107)	(0.0112)	(0.0112)	(0.0106)	(0.0106)	(0.0107)
female	-0.0114	-0.00649	-0.0239***	-0.0505***	-0.0125	-0.0125	-0.0182**
	(0.00857)	(0.00846)	(0.00889)	(0.00881)	(0.00844)	(0.00844)	(0.00847)
wealth Q2	0.0172	0.0339***	0.122***	-0.00787	0.0393***	0.0393***	0.0386***
	(0.0123)	(0.0121)	(0.0128)	(0.0128)	(0.0121)	(0.0121)	(0.0121)
wealth Q3	0.0891***	0.124***	0.193***	0.0843***	0.136***	0.136***	0.145***
	(0.0125)	(0.0123)	(0.0130)	(0.0129)	(0.0123)	(0.0123)	(0.0123)
wealth Q4	0.166***	0.203***	0.268***	0.177***	0.234***	0.234***	0.248***
	(0.0127)	(0.0125)	(0.0132)	(0.0131)	(0.0125)	(0.0125)	(0.0125)
wealth Q5	0.236***	0.260***	0.322***	0.255***	0.310***	0.310***	0.337***
	(0.0130)	(0.0128)	(0.0135)	(0.0134)	(0.0128)	(0.0128)	(0.0129)
N	46153	46158	46022	44881	46142	46142	46128
adj. R^2	0.031	0.039	0.045	0.039	0.045	0.045	0.049
F	130.7	164.3	195.7	162.4	196.2	196.2	211.8

All regressions restrict the HRS respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. All the regressions include cognitive score, three age dummies, indicators for phone-interview mode, high school graduation, female, and wealth quantile as a set of controls. Also, all regressions include survey-year fixed effects. The dependent variables are the proxy *Moon*, but one question is taken away. The taken-out question is shown in the column label. All the proxies are standardized. The respondents aged between 50 and 89 serve as the base level. Standard errors are in parentheses.

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Table 5: Robustness Check II: Adding more questions into the proxies

	(1)	(2)	(3)	(4)
	without individual fixed effects		with individual fixed effects	
cognitive score	0.0995***	0.106***	0.0530***	0.0547***
	(0.00548)	(0.00596)	(0.00837)	(0.00934)
age	-0.0128***	-0.0110***	-0.0167*	-0.0271***
	(0.000731)	(0.000796)	(0.00972)	(0.0105)
phone	-0.0224**	-0.0202*	-0.0141	-0.0109
	(0.00968)	(0.0105)	(0.00902)	(0.0102)
high school	-0.0309**	-0.000113		
	(0.0123)	(0.0134)		
female	-0.103***	-0.0330***		
	(0.0105)	(0.0114)		
wealth Q2	-0.0265*	0.0313*	0.0560***	0.0299
	(0.0153)	(0.0167)	(0.0215)	(0.0248)
wealth Q3	0.159***	0.132***	0.261***	0.117***
	(0.0155)	(0.0169)	(0.0242)	(0.0266)
wealth Q4	0.196***	0.246***	0.300***	0.192***
	(0.0153)	(0.0166)	(0.0265)	(0.0291)
wealth Q5	0.275***	0.324***	0.426***	0.312***
	(0.0153)	(0.0166)	(0.0291)	(0.0318)
N	27648	27614	25017	24985
adj. R^2	0.064	0.057	0.349	0.303
F	205.5	182.6	48.68	24.61

All regressions restrict the HRS respondents aged between 65 and 89 who participated in the survey from 2004 to 2018. All the regressions include cognitive score, age, indicators for phone-interview mode, high school graduation, female, and wealth quantile as a set of controls. Also, all regressions include survey-year fixed effects. The dependent variables are based on the proxy *Moon*, but columns 1 and 3 add the social security income and vehicle-value questions into the proxies. Columns 2 and 4 use the original set of questions to construct the proxy for a comparison. All the proxies are standardized. Standard errors are clustered at the individual level for columns 3 and 4. Standard errors are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

Table 6: Descriptive statistics for the selected questions in the PSID

variable	maximal rounding (%)		Opt-out (%)	
		Don't know	Refused to Answer	skip
Home value	30.5	3.0	1.7	0.5
Wage	20.2	1.5	2.7	14.1
Rent	28.8	0.5	3.2	0.7
Property tax	20.7	4.2	0.9	2.9
Food home	66.1	1.9	1.4	
Food out	71.7	1.8	1.1	

Each row summarizes the response format. The maximal rounding indicates the numerical response format in which the number is rounded to the precision of 1. That is, the leftmost digit is any number, and the rest of the digit positions are zeros, for example, 2,000 or 100,000. The respondents in the PSID can opt out of the questions, choosing *Don't know, Refused to answer*, or skipping the question. The column *skip* counts the missing responses only for those who have access to the question but skip it. The access is determined by the following two questions: *Do (you (or anyone else in your family living there) / they (or anyone else in the family living there)) own the (apartment/mobile home/home), pay rent, or what?* and *(Are/is) (you/he/she) working now, looking for work, retired, a student, keeping house, or what?*. The first question determines whether one has access to the question of home value, rent, and property tax. The latter question indicates the employment status and determines the access to the wage question. For example, if one answers *working now* to the latter question but there is no response on the wage question, I count it as *skip*. There are no indicative questions about access to food consumption. The sample is limited to the household heads aged between 50 and 89 who participated in the survey from 1999 to 2019.

The full question texts are:
Home value: *Could you tell me what the present value of your (house/apartment) is–I mean about how much would it bring if you sold it today?*
Wage: *How much did you (HEAD) earn altogether from wages or salaries, that is, before anything was deducted for taxes or other things?*
Rent: *About how much rent do you pay a month?–AMOUNT*
Property tax: *About how much are your total yearly property taxes, including city, county, and school taxes?*
Food home: *How much do you spend on that food in an average week?–AMOUNT*
Food out: *About how much do you spend eating out?–AMOUNT*

Table 7: Regression analysis of the proxies from the PSID

	(1)	(2)	(3)
age	-0.00741***		
	(0.000593)		
$60 \leq$ age < 70		-0.0317**	-0.0405***
		(0.0133)	(0.0136)
$70 \leq$ age < 80		-0.121***	-0.144***
		(0.0165)	(0.0170)
$80 \leq$ age		-0.268***	-0.307***
		(0.0223)	(0.0229)
high school	0.213***	0.214***	0.130***
	(0.0170)	(0.0169)	(0.0183)
female	-0.0311***	-0.0305***	-0.00116
	(0.0114)	(0.0114)	(0.0117)
black			-0.304***
			(0.0179)
wealth Q2	0.0936***	0.0915***	0.0801***
	(0.0258)	(0.0258)	(0.0270)
wealth Q3	-0.00706	-0.00891	-0.0471**
	(0.0228)	(0.0228)	(0.0236)
wealth Q4	-0.00831	-0.0123	-0.0994***
	(0.0206)	(0.0206)	(0.0218)
wealth Q5	0.242***	0.234***	0.0461**
	(0.0199)	(0.0198)	(0.0221)
mean (proxy)	.0421	.0421	.0542
N	33424	33424	30939
adj. R^2	0.027	0.028	0.060
F	131.6	104.2	82.77

All regressions restrict the PSID respondents aged between 50 and 89 who participated in the survey from 1999 to 2019. The dependent variables are the proxy *Moon* using six questions (see the list of questions in Table 6.). All the regressions include indicators for high school graduation, female, and wealth quantile as a set of controls, and all the regressions include the survey-year fixed effects. Column 1 controls for age as a continuous variable. Motivated by Figure 7, columns 2–3 include three age dummies instead of adding the age variable. Each dummy spans ten years, and the respondents aged between 50 and 89 serve as the base level. Columns 1 and 2 follow a similar specification as the HRS analysis. Column 3 employs demographic variables only available in the PSID. The regression in column 3 includes the indicator of African Americans. It also includes the grown-up state and rural-area indicator fixed effects. The regression analysis for the other proxies is in Appendix A.3. Standard errors are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

Table 8: Descriptive statistics for PSID respondents in the South

	(1)	(2)	(3)	(4)	(5)	(6)
	before	after	before	after	before	after
	overall		rural		urban	
age	48.97	38.55	49.08	38.46	48.93	38.59
black	0.73	0.65	0.77	0.71	0.72	0.62
female	0.61	0.62	0.58	0.62	0.62	0.61
year of schooling	12.7	13.1	12.2	12.9	12.9	13.2
rural	0.27	0.3				
observations	7,582	7,896	2,064	2,408	5,518	5,488

A pooled sample of the PSID household heads represents 15,478 individual-years, 2,738 unique individuals who participated in the survey from 1999 to 2019. The sample is limited to those born between 1955 and 1975 in the Southern states. Columns 1 and 2 present the average demographic variables for this sample, disaggregated by the enactment of the integration policy. *before* and *after* indicate that the sample is limited to the respondents born before and after 1965. Columns 3–6 further disaggregate these samples by whether the grown-up area is urban or rural. Each row summarizes the corresponding average.

Table 9: Impact of the integration policy on Cognitive proxy in the South

	(1) rural black	(2) urban black	(3) rural white	(4) urban white
$50 \leq$ age < 60	0.0352	-0.106	-0.0613	-0.0633
	(0.0756)	(0.107)	(0.0479)	(0.0584)
$60 \leq$ age	0.0428	-0.282	-0.192*	0.0990
	(0.146)	(0.262)	(0.107)	(0.117)
post-1965 trend break	0.0297***	0.00609	0.00313	-0.00476
	(0.0113)	(0.0129)	(0.00619)	(0.00780)
high school	0.236***	-0.0183	0.260***	0.218**
	(0.0816)	(0.109)	(0.0562)	(0.101)
female	0.0438	-0.0767	-0.0212	0.0104
	(0.0675)	(0.0716)	(0.0395)	(0.0433)
wealth Q2	0.274***	-0.286***	0.167***	0.0229
	(0.0639)	(0.110)	(0.0392)	(0.0691)
wealth Q3	0.122*	-0.201*	0.0634	0.0534
	(0.0664)	(0.107)	(0.0415)	(0.0623)
wealth Q4	-0.0274	-0.0160	-0.00306	0.0334
	(0.0785)	(0.124)	(0.0480)	(0.0664)
wealth Q5	-0.152	-0.0225	0.0217	0.116*
	(0.124)	(0.116)	(0.0654)	(0.0658)
N	3205	1119	7248	3518
adj. R^2	0.027	0.108	0.021	0.025
F	4.345	2.342	5.983	1.488

The estimates are based on Eq. (3). The dependent variables are the proxy *Moon*. All regressions restrict the PSID household heads born between 1955 and 1975 in Southern states who participated in the 1999–2019 surveys. Columns 1 and 2 use the samples further limited to the black respondents who grew up in a rural area and urban area, respectively. Columns 3 and 4 use the white respondents who grew up in a rural area and urban area, respectively. All the columns include age dummies, the trend break interaction term, indicators for high school graduation, female, and wealth as a set of controls. Also, all the regressions include the survey-year fixed effects, grown-up state fixed effects, and the grown-up rural indicator fixed effects. The respondents aged below 50 serve as the base level. Standard errors are clustered at the individual level and are in parentheses.

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Table 10: Triple Difference estimate of the Impact of 1964 Civil Rights Act

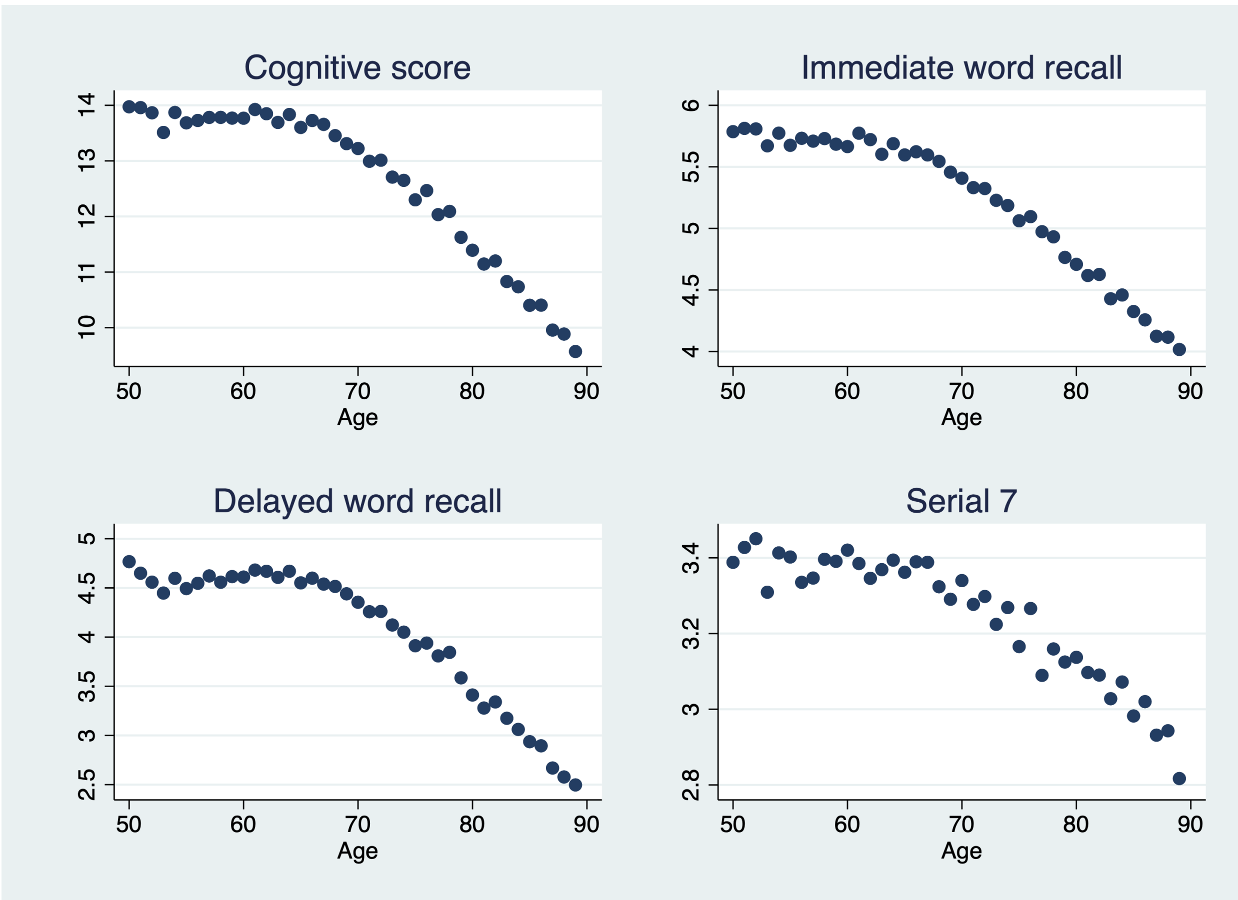
	(1) Moon	(2) Diabetes	(3) Ear	(4) High school
post-1965 $\times$ black $\times$ rural	0.266**	-0.0197**	-0.156***	-0.0124
	(0.119)	(0.00944)	(0.0562)	(0.0804)
post-1965 $\times$ rural	-0.268***	0.00893	0.108**	0.111
	(0.0941)	(0.00572)	(0.0514)	(0.0678)
post-1965 $\times$ black	0.0844	0.0000971	0.0216	0.0218
	(0.0584)	(0.000653)	(0.0341)	(0.0312)
rural $\times$ black	-0.232***	0.00941	0.0904**	0.0584
	(0.0887)	(0.00760)	(0.0377)	(0.0689)
post-1965	-0.0193	-0.000128	0.00547	-0.0164
	(0.0472)	(0.000964)	(0.0307)	(0.0216)
rural	0.146**	0.00243*	-0.0611*	-0.123**
	(0.0714)	(0.00125)	(0.0330)	(0.0594)
black	-0.240***	-0.000893	-0.0778***	-0.0547***
	(0.0476)	(0.000821)	(0.0261)	(0.0211)
N	15091	14855	14855	15469
adj. R^2	0.030	0.015	0.091	0.057
F	9.833	0.449	2.406	2.357

Estimates are based on Eq. (4). All regressions restrict the PSID household heads born between 1955 and 1975 in Southern states who participated in the 1999–2019 surveys. The estimation employs three variations: policy, race, and location. *post-1965* captures whether one is born after the integration policy. *black* indicates individual race, and *rural* denotes whether one grew up rural area. All the columns include age dummies, indicators for high school graduation, female, and wealth quantile as a set of controls, but the coefficients on the controls are suppressed. Also, all the regressions include the survey-year fixed effects and grown-up state fixed effects. The first column uses the proxy *Moon* as a robustness check for the findings in Table 9. Columns 2 and 3 use health outcomes before the respondents grow older. The dependent variable in column 2 is whether one had diabetes before 17 years old. Column 3 asks whether one had chronic ear problems before 17 years old. These questions are more direct to the fetal origin hypothesis because asking the same questions when they fully mature could be biased by many other factors. The last column uses the indicator for high school graduation as a dependent variable. Standard errors are clustered at the individual level and are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

Figure 1: Response heaping

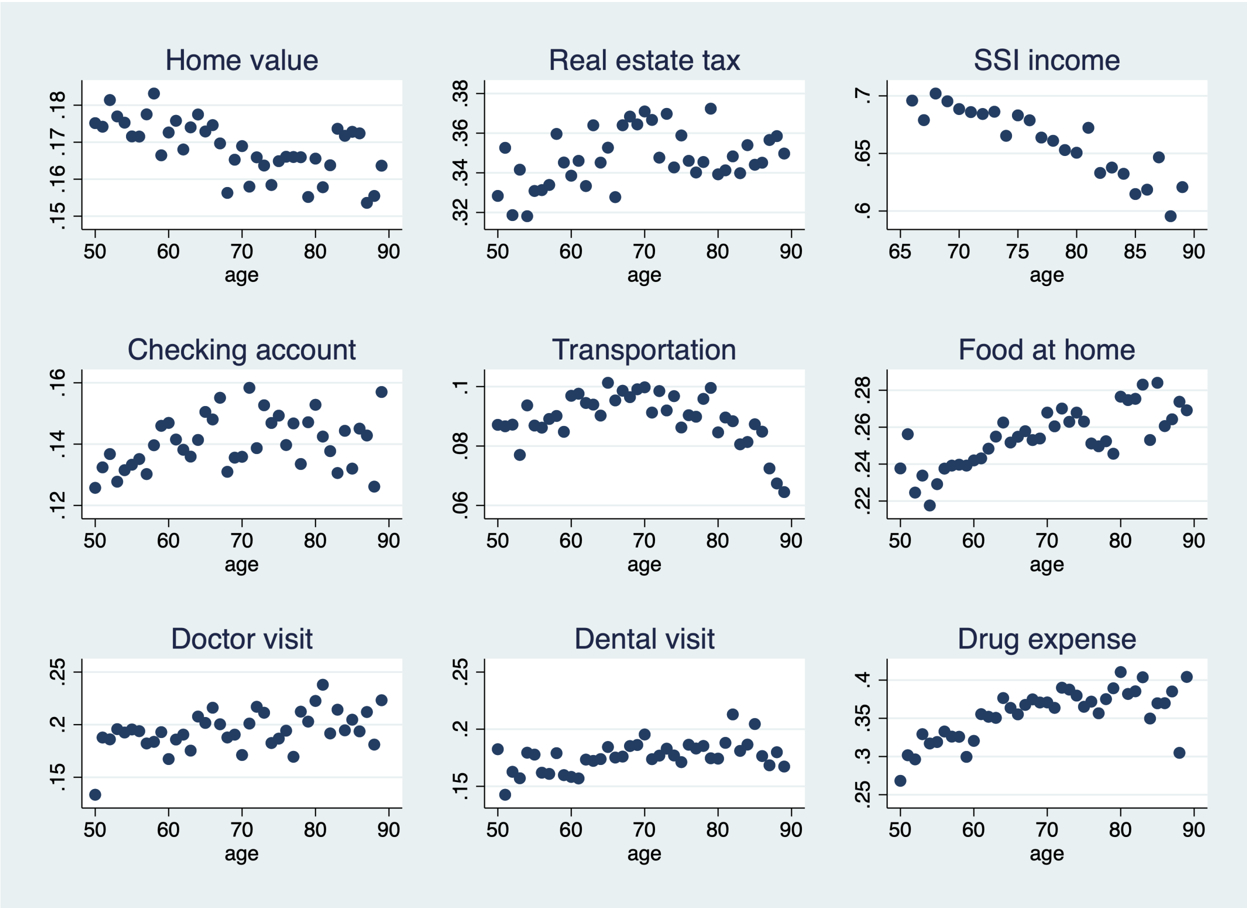
The figure plots the numerical responses of the selected questions from the HRS. All the figures restrict the household head respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. The vertical axis represents frequency indicating the number of responses on a specific numerical value denoted by the horizontal axis. See the full description of the questions in Table 2.

Figure 2: Age trend of cognitive scores



The figure plots the average score of the cognition-related questions against age in years available in the HRS. All the figures restrict the household head respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. The first panel plots the sum of the three questions: immediate word recall, delayed word recall, and serial 7's test. The rest of the panels present the score of each question against age in years.

Figure 3: The level of round by age

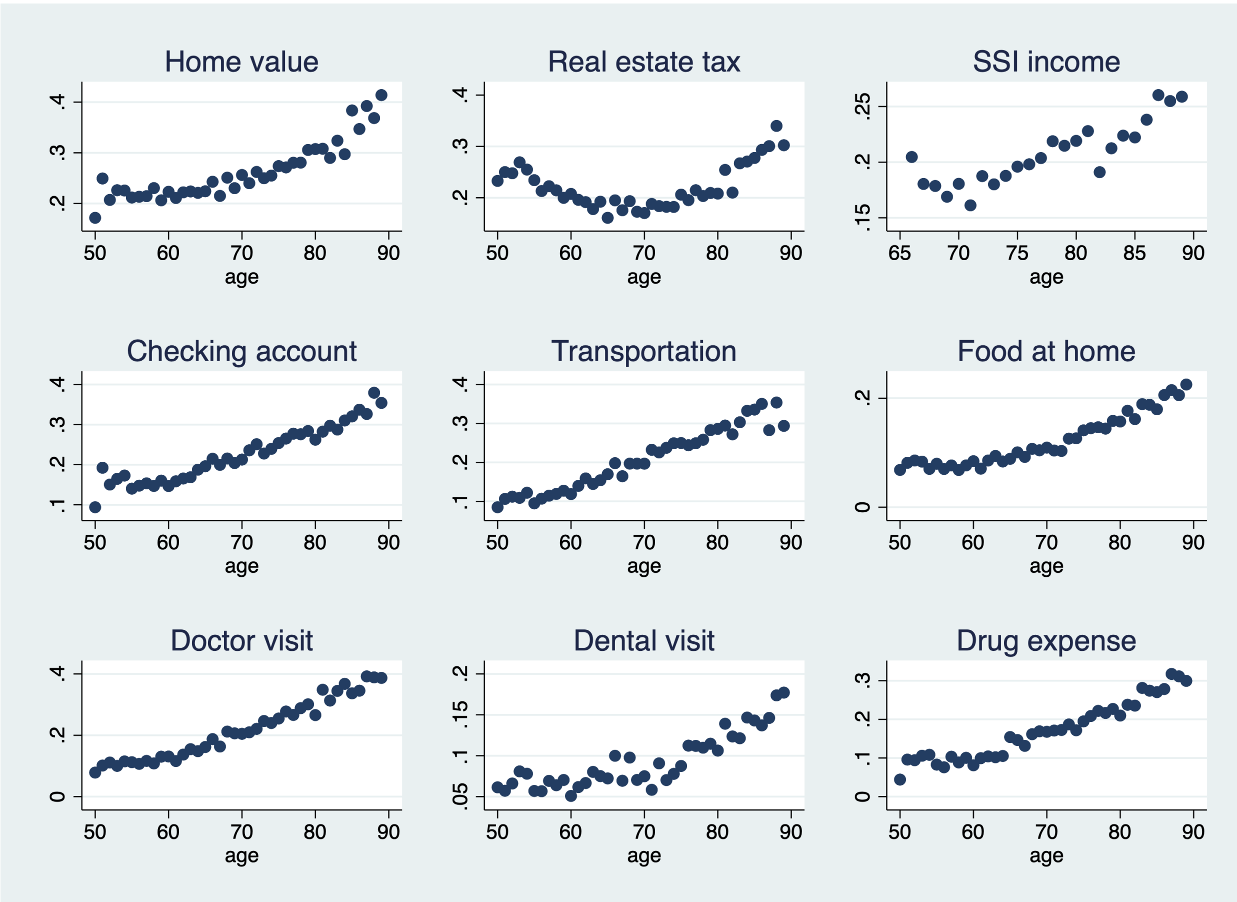


All the panels restrict the household head respondents aged between 50 and 89 who participated in the HRS survey from 2004 to 2018, except for the social security income panel. Each panel summarizes the level of round on the numerical response against age in years. Note that Gideon, Helppie-McFall, and Hsu (2017) define the level of rounding as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. The higher the level of round, the more the respondents round up the responses. The panel title indicates the question.

Figure 4: Maximal rounding by age

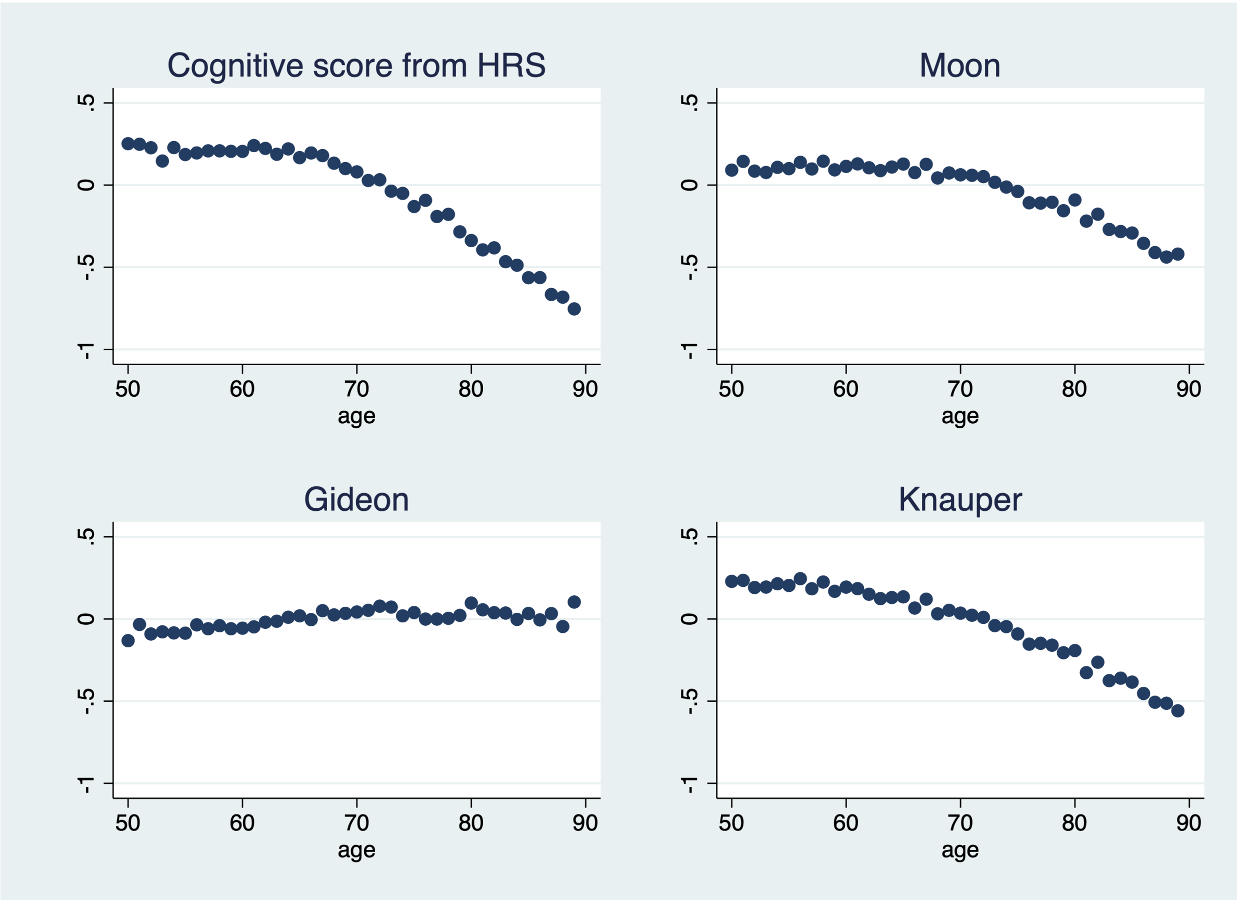
All the panels restrict the household head respondents aged between 50 and 89 who participated in the HRS survey from 2004 to 2018, except for the social security income panel. Each panel summarizes whether one uses the maximal rounding conditioning that the respondent provides a numerical response. The maximal rounding indicates the numerical response format in which the number is rounded to the precision of 1. That is, the leftmost digit is any number, and the rest of the digit positions are zeros, for example, 2,000 or 100,000. Each panel represents the share of the maximal rounding out of numerical responses by age. The panel title indicates the question.

Figure 5: Opt-out trend by age



All the panels restrict the household head respondents aged between 50 and 89 who participated in the HRS survey from 2004 to 2018, except for the social security income panel. Each panel summarizes whether the respondents choose the opt-out responses. The HRS provides the opt-out options for the open-ended numerical questions, such as for *Don't know* and *Refused to answer*. Each panel represents the share of opt-out responses by age. The panel title indicates the question.

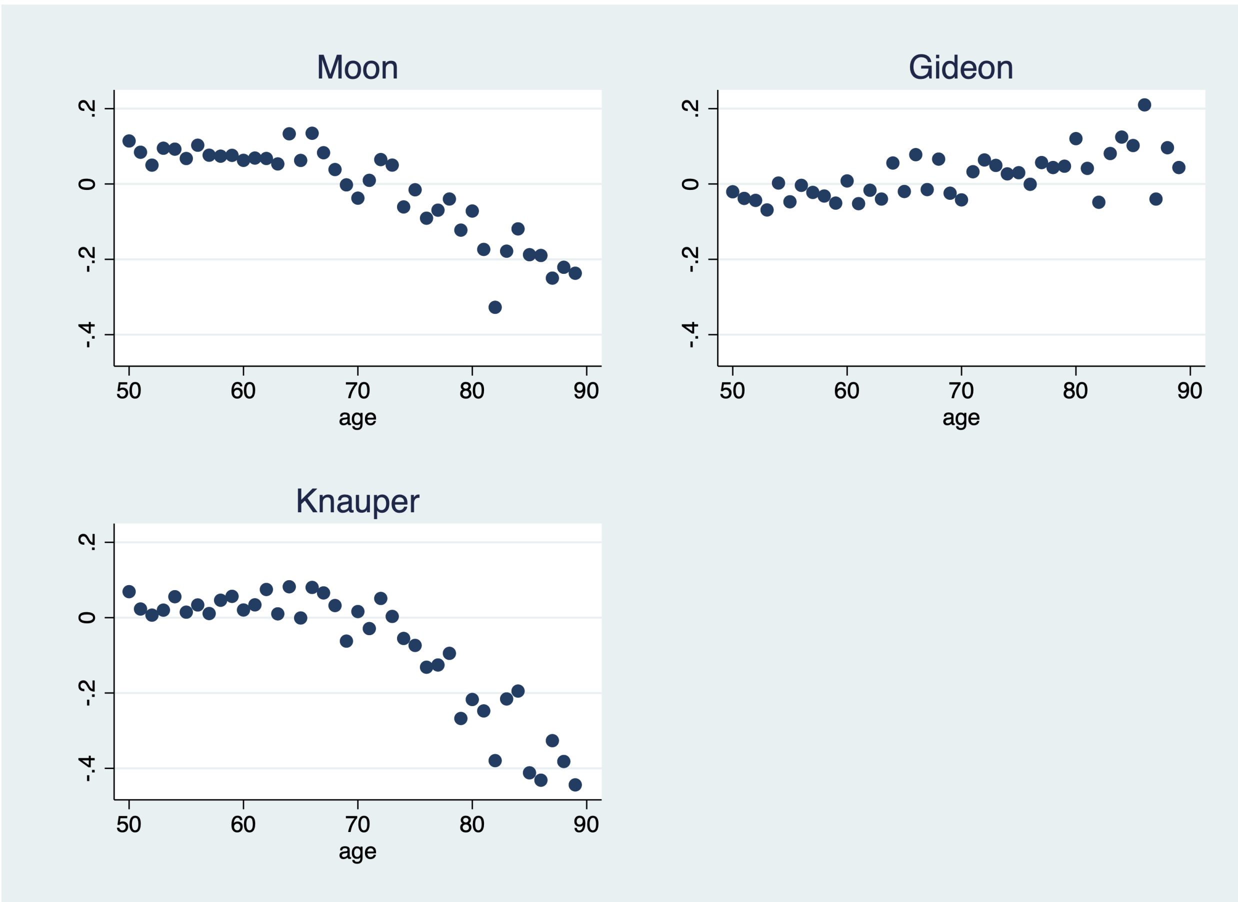
Figure 6: Age trend of the cognitive proxies



The figure plots the average cognitive proxies against age in years, constructed from the HRS data. All the panels are limited to the household heads aged between 50 and 89 who participated in the survey from 2004 to 2018. The first panel is directly from Figure 2 and serves as a benchmark. The rest of the panels present the average of the proxies against age in years. The cognitive score and all the proxies are standardized.

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it. For more details, see Table 2.

Figure 7: Age trend of the cognitive proxies in the PSID



The figure plots the average cognitive proxies against age in years, constructed from the PSID data. All the figures restrict the PSID household heads aged between 50 and 89 who participated in the survey from 1999 to 2019. Each panel presents the average of the proxies against age in years. All the proxies are standardized.

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits}-\text{the number of significant digits}}{\text{the number of total digits}-1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know* or *Refused to answer*. For more details, see Table 6.

References

Almond, Douglas, Kenneth Y Chay, and Michael Greenstone. 2006. "Civil rights, the war on poverty, and black-white convergence in infant mortality in the rural South and Mississippi."

Almond, Douglas, and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis." *Journal of economic perspectives* 25 (3): 153–72.

Andrews, Frank M, and A Regula Herzog. 1986. "The quality of survey data as related to age of respondent." *Journal of the American Statistical Association* 81 (394): 403–410.

Barbosa-Silva, Maria Carolina, Luís Eduardo Santos, and Bárbara Rangel. 2018. "The impact of non-neurotropic influenza strains on the brain: a role for microglial priming?" *Journal of Neuroscience* 38 (36): 7758–7760.

Boyle, Phelim P, and Cormac Ó Gráda. 1986. "Fertility trends, excess mortality, and the Great Irish Famine." *Demography,* 543–562.

Brown, M Christopher, et al. 1999. *The quest to define collegiate desegregation: Black colleges, Title VI compliance, and post-Adams litigation.* ERIC.

Bucher-Koenen, Tabea, and Michael Ziegelmeyer. 2011. "Who lost the most? Financial literacy, cognitive abilities, and the financial crisis."

Budd, John W, and Timothy Guinnane. 1991. "Intentional age-misreporting, age-heaping, and the 1908 Old Age Pensions Act in Ireland." *Population Studies* 45 (3): 497–518.

Cagney, Kathleen A, and Diane S Lauderdale. 2002. "Education, wealth, and cognitive function in later life." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 57 (2): P163–P172.

Chandra, Amitabh, Courtney Coile, and Corina Mommaerts. Forthcoming. "What Can Economics Say about Alzheimer's Disease?" *Journal of Economic Literature.*

Cole, Shawn A, and Gauri Kartini Shastry. 2008. *If you are so smart, why aren't you rich?: the effects of education, financial literacy and cognitive ability on financial market participation.* Citeseer.

Colsher, Patricia L, and Robert B Wallace. 1989. "Data quality and age: health and psychobehavioral correlates of item nonresponse and inconsistent responses." *Journal of Gerontology* 44 (2): P45–P52.

Cutler, David M, Adriana Lleras-Muney, and Tom Vogl. 2008. "Socioeconomic status and health: dimensions and mechanisms."

De Boo, Hendrina A, and Jane E Harding. 2006. "The developmental origins of adult disease (Barker) hypothesis." *Australian and New Zealand Journal of Obstetrics and Gynaecology* 46 (1): 4–14.

Gibson, Allison K, and Virginia E Richardson. 2017. "Living alone with cognitive impairment: Findings from the National Health and Aging Trends Study." *American Journal of Alzheimer's Disease & Other Dementias®* 32 (1): 56–62.

Gideon, Michael, Brooke Helppie-McFall, and Joanne W Hsu. 2017. "Heaping at round numbers on financial questions: The role of satisficing." In *Survey research methods,* 11:189. 2. NIH Public Access.

Herzog, A Regula, and Robert B Wallace. 1997. "Measures of cognitive functioning in the AHEAD Study." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 52 (Special_Issue): 37–48.

Holbrook, Allyson L, Sowmya Anand, Timothy P Johnson, Young Ik Cho, Sharon Shavitt, Noel Chávez, and Saul Weiner. 2014. "Response heaping in interviewer-administered surveys: Is it really a form of satisficing?" *Public Opinion Quarterly* 78 (3): 591–633.

Jessen, Frank, Rebecca E Amariglio, Martin Van Boxtel, Monique Breteler, Mathieu Ceccaldi, Gaël Chételat, Bruno Dubois, Carole Dufouil, Kathryn A Ellis, Wiesje M Van Der Flier, et al. 2014. "A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease." *Alzheimer's & dementia* 10 (6): 844–852.

Karr, Justin E, Raquel B Graham, Scott M Hofer, and Graciela Muniz-Terrera. 2018. "When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death." *Psychology and aging* 33 (2): 195.

Kaufman, Alan S, and John L Horn. 1996. "Age changes on tests of fluid and crystallized ability for women and men on the Kaufman Adolescent and Adult Intelligence Test (KAIT) at ages 17–94 years." *Archives of clinical neuropsychology* 11 (2): 97–121.

Knäuper, Barbel, Robert F Belli, Daniel H Hill, and A Regula Herzog. 1997. "Question difficulty and respondents' cognitive ability: The effect on data quality." *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-* 13:181–199.

Krosnick, Jon A. 1991. "Response strategies for coping with the cognitive demands of attitude measures in surveys." *Applied cognitive psychology* 5 (3): 213–236.

Manski, Charles F, and Francesca Molinari. 2010. "Rounding probabilistic expectations in surveys." *Journal of Business & Economic Statistics* 28 (2): 219–231.

Mayeda, Elizabeth Rose, Taylor M. Mobley, Robert E. Weiss, Audrey R. Murchland, Lisa F. Berkman, and Erika L. Sabbath. 2020. "Association of work-family experience with mid- and late-life memory decline in US women." *Neurology* 95 (23): e3072–e3080.

Muñoz-Murillo, Melisa, Pilar B Álvarez-Franco, and Diego A Restrepo-Tobón. 2020. "The role of cognitive abilities on financial literacy: New experimental evidence." *Journal of Behavioral and Experimental Economics* 84:101482.

Muralidharan, Karthik, and Nishith Prakash. 2017. "Cycling to school: Increasing secondary school enrollment for girls in India." *American Economic Journal: Applied Economics* 9 (3): 321–50.

Murman, Daniel L. 2015. "The impact of age on cognition." In *Seminars in hearing,* 36:111–121. 03. Thieme Medical Publishers.

Myers, Robert J. 1976. "An instance of reverse heaping of ages." *Demography* 13 (4): 577–580.

Nelson, Alan. 2002. "Unequal treatment: confronting racial and ethnic disparities in health care." *Journal of the national medical association* 94 (8): 666.

Persson, Petra, and Maya Rossin-Slater. 2018. "Family ruptures, stress, and the mental health of the next generation." *American economic review* 108 (4-5): 1214–52.

Portacolone, Elena, Julene K Johnson, Kenneth E Covinsky, Jodi Halpern, and Robert L Rubinstein. 2018. "The effects and meanings of receiving a diagnosis of mild cognitive impairment or Alzheimer's disease when one lives alone." *Journal of Alzheimer's Disease* 61 (4): 1517–1529.

Riddles, Minsun K, Sharon L Lohr, J Michael Brick, Patrick T Langetieg, John M Payne, and Alan H Plumley. 2017. "Handling respondent rounding of wages using the IRS and CPS matched dataset." *2016 IRS Research Bulletin,* 60–68.

Roberts, John M, and Devon D Brewer. 2001. "Measures and tests of heaping in discrete quantitative distributions." *Journal of Applied Statistics* 28 (7): 887–896.

Rushton, J Philippe, and Arthur R Jensen. 2005. "Thirty years of research on race differences in cognitive ability." *Psychology, public policy, and law* 11 (2): 235.

Salthouse, Timothy A. 2010. "Selective review of cognitive aging." *Journal of the International neuropsychological Society* 16 (5): 754–760.

Schady, Norbert, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, Karen Macours, Daniela Marshall, Christina Paxson, et al. 2015. "Wealth gradients in early childhood cognitive development in five Latin American countries." *Journal of Human Resources* 50 (2): 446–463.

Sonnega, Amanda, Jessica D Faul, Mary Beth Ofstedal, Kenneth M Langa, John WR Phillips, and David R Weir. 2014. "Cohort profile: the health and retirement study (HRS)." *International journal of epidemiology* 43 (2): 576–585.

Weir, David R. 2017. "Health and Retirement Study Imputation of Cognitive Functioning Measures."

Zinn, S, and A Würbach. 2016. "A statistical approach to address the problem of heaping in self-reported income data." *Journal of Applied Statistics* 43 (4): 682–703.

A Appendix

Table A.1: Regression analysis on the proxies with more age groups

	(1) Moon	(2) Gideon	(3) Knäuper
$55 \leq$ age < 60	0.00261	0.0203	-0.00999
	(0.0159)	(0.0174)	(0.0141)
$60 \leq$ age < 65	0.00916	0.0656***	-0.0324**
	(0.0161)	(0.0176)	(0.0142)
$65 \leq$ age < 70	0.00441	0.106***	-0.0811***
	(0.0162)	(0.0177)	(0.0143)
$70 \leq$ age < 75	-0.0112	0.129***	-0.123***
	(0.0163)	(0.0178)	(0.0144)
$75 \leq$ age < 80	-0.0933***	0.125***	-0.213***
	(0.0168)	(0.0183)	(0.0148)
$80 \leq$ age < 85	-0.146***	0.145***	-0.304***
	(0.0179)	(0.0195)	(0.0158)
$85 \leq$ age	-0.222***	0.127***	-0.376***
	(0.0205)	(0.0224)	(0.0180)
cognitive score	0.0896***	0.0110**	0.0728***
	(0.00450)	(0.00492)	(0.00397)
phone	-0.0112	-0.0193**	-0.00118
	(0.00783)	(0.00855)	(0.00691)
high school	0.0118	-0.0736***	0.0312***
	(0.0106)	(0.0116)	(0.00936)
female	-0.0195**	0.0482***	-0.0448***
	(0.00841)	(0.00919)	(0.00742)
N	46162	45841	46162
adj. R^2	0.048	0.016	0.064
F	152.8	46.10	196.8

All regressions restrict the HRS respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. All the regressions include cognitive score, indicators for phone-interview mode, high school graduation, female, and wealth quantile as a set of controls. Also, all regressions include survey-year fixed effects. The dependent variables are the cognitive proxies defined below. All the proxies are standardized. Instead of adding age terms, all specifications include seven age dummies. Each dummy spans five years, and the first seven rows indicate the corresponding age bins. The age between 50 and 55 serves as the base level. Standard errors are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it. For more details, see Table 2.

Table A.2: Regression analysis on the proxies with individual fixed effects

	(1) Moon	(2) Gideon	(3) Knäuper
cognitive score	0.0436***	0.00418	0.0394***
	(0.00689)	(0.00773)	(0.00591)
age	-0.0200**	-0.0240***	-0.00363
	(0.00811)	(0.00898)	(0.00598)
phone	-0.00554	-0.0241***	0.0105
	(0.00765)	(0.00866)	(0.00642)
wealth Q2	0.0354**	-0.0427**	0.0490***
	(0.0176)	(0.0206)	(0.0157)
wealth Q3	0.109***	-0.0576***	0.112***
	(0.0188)	(0.0216)	(0.0168)
wealth Q4	0.175***	-0.0801***	0.194***
	(0.0208)	(0.0231)	(0.0187)
wealth Q5	0.291***	-0.0711***	0.300***
	(0.0232)	(0.0250)	(0.0207)
mean (proxy)	42893	42557	42893
N	42893	42557	42893
adj. R^2	0.274	0.181	0.329
F	35.50	3.829	43.20

All regressions restrict the HRS respondents aged between 50 and 89 who participated in the survey from 2004 to 2018. All the regressions include cognitive score, age, indicators for phone-interview mode, and wealth quantile as a set of controls. Also, all regressions include the survey-year fixed effects and the individual fixed effects. The dependent variables are the cognitive proxies defined below. All the proxies are standardized. Standard errors are clustered at the individual level and are in parentheses.

* $p < 0.10$, ** $p < .05$, *** $p < .01$

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it. For more details, see Table 2.

Table A.3: Regression analysis of the proxies from the PSID

	(1) Gideon	(2) Gideon	(3) Gideon	(4) Knäuper	(5) Knäuper	(6) Knäuper
age	0.00672***			-0.00788***		
	(0.000578)			(0.000617)		
$60 \leq$ age < 70		0.0997***	0.103***		-0.0171	-0.0265*
		(0.0129)	(0.0135)		(0.0138)	(0.0143)
$70 \leq$ age < 80		0.136***	0.130***		-0.0933***	-0.117***
		(0.0161)	(0.0168)		(0.0172)	(0.0178)
$80 \leq$ age		0.177***	0.166***		-0.320***	-0.348***
		(0.0218)	(0.0228)		(0.0232)	(0.0239)
high school	-0.00788	-0.0124	-0.0635***	0.139***	0.142***	0.0770***
	(0.0166)	(0.0165)	(0.0182)	(0.0177)	(0.0176)	(0.0192)
female	0.0838***	0.0850***	0.0919***	-0.0576***	-0.0566***	-0.0306**
	(0.0111)	(0.0111)	(0.0115)	(0.0119)	(0.0118)	(0.0122)
black			-0.0521***			-0.265***
			(0.0177)			(0.0187)
wealth Q2	0.137***	0.137***	0.155***	0.0191	0.0159	-0.0197
	(0.0250)	(0.0250)	(0.0267)	(0.0268)	(0.0268)	(0.0282)
wealth Q3	-0.250***	-0.250***	-0.265***	-0.215***	-0.217***	-0.261***
	(0.0222)	(0.0222)	(0.0234)	(0.0237)	(0.0237)	(0.0247)
wealth Q4	-0.484***	-0.482***	-0.513***	-0.244***	-0.250***	-0.330***
	(0.0201)	(0.0201)	(0.0216)	(0.0215)	(0.0214)	(0.0227)
wealth Q5	-0.538***	-0.536***	-0.584***	0.0203	0.00855	-0.180***
	(0.0193)	(0.0193)	(0.0219)	(0.0207)	(0.0207)	(0.0231)
N	33107	33107	30641	33424	33424	30939
adj. R^2	0.056	0.056	0.065	0.024	0.025	0.048
F	276.9	214.7	186.2	109.6	90.12	78.82

All regressions restrict the PSID respondents aged between 50 and 89 who participated in the survey from 1999 to 2019. The dependent variables are the proxy *Gideon* and *Knäuper* using 6 questions (see the list of questions in Table 6.). All the regressions include indicators for high school graduation, female and wealth quantile as a set of controls, and all the regressions include survey-year fixed effects. Columns 1 and 4 control for age as a continuous variable. Motivated by Figure 7, columns 2–3 and 5–6 include three age dummies instead of adding the age variable. Each dummy spans ten years, and the respondents aged between 50 and 59 serve as the base level. Columns 1, 2, 4, and 5 follow a similar specification as the HRS analysis. Columns 3 and 6 employ demographic variables only available in the PSID. The regressions in columns 3 and 6 include the indicator of African Americans. These columns also include grown-up state and rural-area indicator fixed effects. Standard errors are in parentheses.

* $p < 0.10$, ** $p < .05$, *** $p < .01$

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know* or *Refused to answer*. For more details, see Table 6.

Table A.4: Regression analysis on the proxies with individual fixed effects in the PSID

	(1) Moon	(2) Gideon	(3) Knäuper
age	-0.00669***	-0.000484	-0.00142
	(0.00131)	(0.00130)	(0.00120)
wealth Q2	0.0468	0.0524	0.0182
	(0.0339)	(0.0399)	(0.0330)
wealth Q3	-0.0161	-0.0739**	-0.154***
	(0.0332)	(0.0367)	(0.0364)
wealth Q4	-0.118***	-0.181***	-0.331***
	(0.0367)	(0.0373)	(0.0432)
wealth Q5	-0.0959**	-0.219***	-0.305***
	(0.0402)	(0.0408)	(0.0452)
mean (proxy)	.0616	.0104	.0082
N	29739	29434	29739
adj. R^2	0.331	0.346	0.355
F	9.691	9.684	13.51

All regressions restrict the PSID respondents aged between 50 and 89 who participated in the survey from 1999 to 2019. All the regressions include age and wealth quantile as a set of controls. Also, all the regressions include the grown-up state fixed effects and individual fixed effects. The dependent variables are the cognitive proxies defined below. All the proxies are standardized. Standard errors are clustered at the individual level and are in parentheses.
* $p < 0.10$, ** $p < .05$, *** $p < .01$

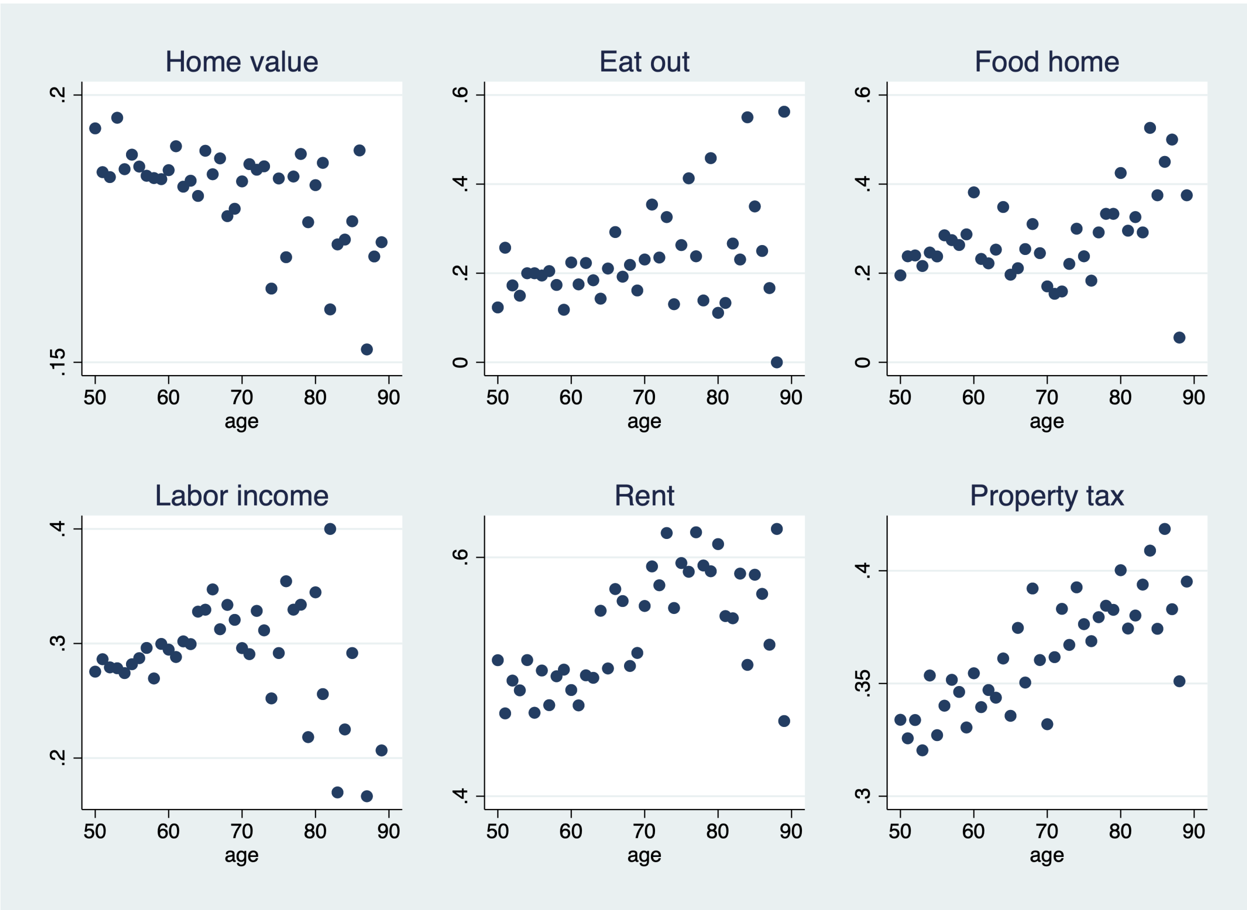
I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know* or *Refused to answer*. For more details, see Table 6.

Figure A.1: Age trend of the cognitive proxies controlling for individual effects

Each proxy is residualized by the non-time varying controls within the same respondents. Then, the corresponding residuals are plotted against age in years. The sample is limited to the household head aged between 50 and 89 who participated in the HRS survey from 2004 to 2018. The first panel serves as a benchmark. The rest of the panels present the average of the residual proxies against age in years. The cognitive score and all proxies are standardized.

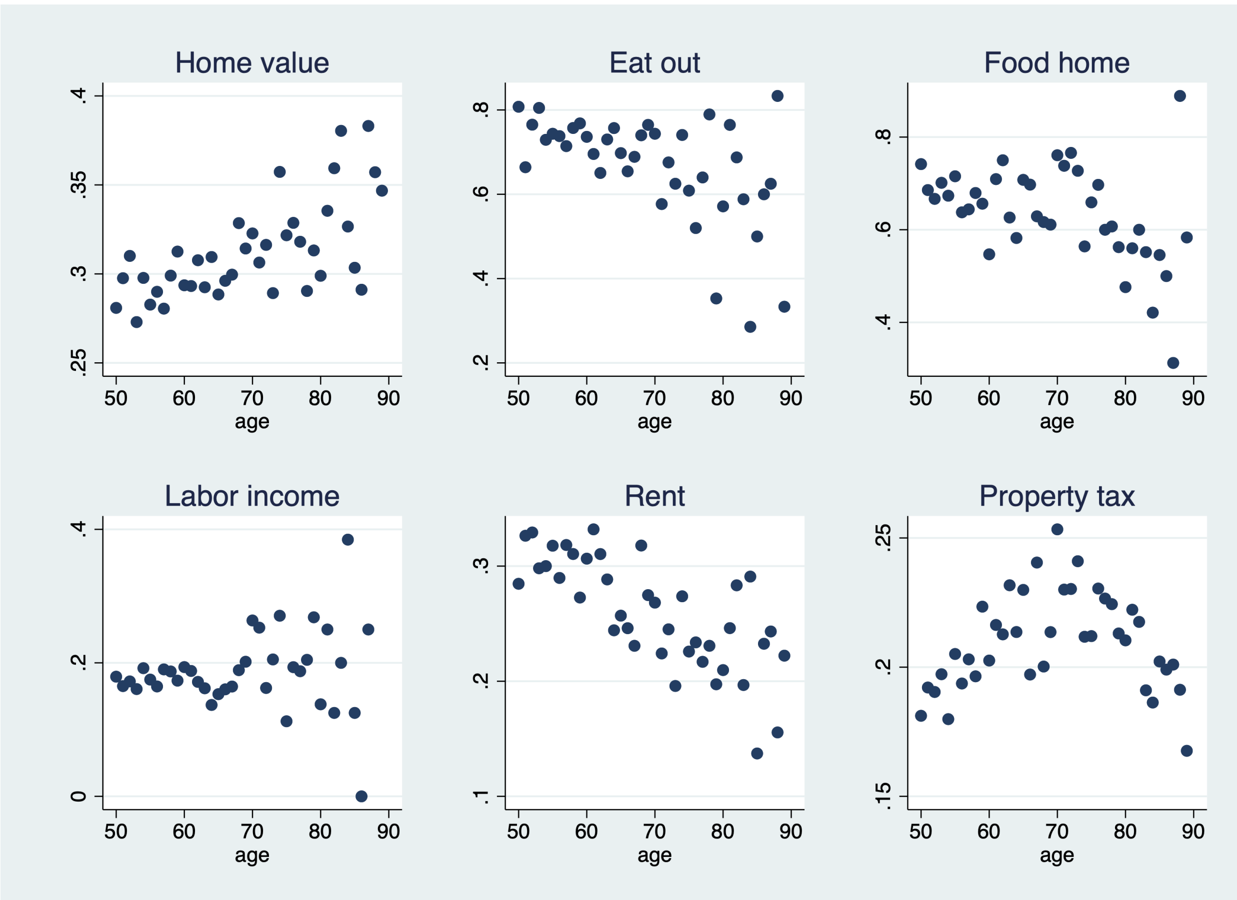
I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know, Refused to answer* or *skip*. *skip* counts the response only for those who have access to the question but skip it. For more details, see Table 2.

Figure A.2: The level of round by age in PSID



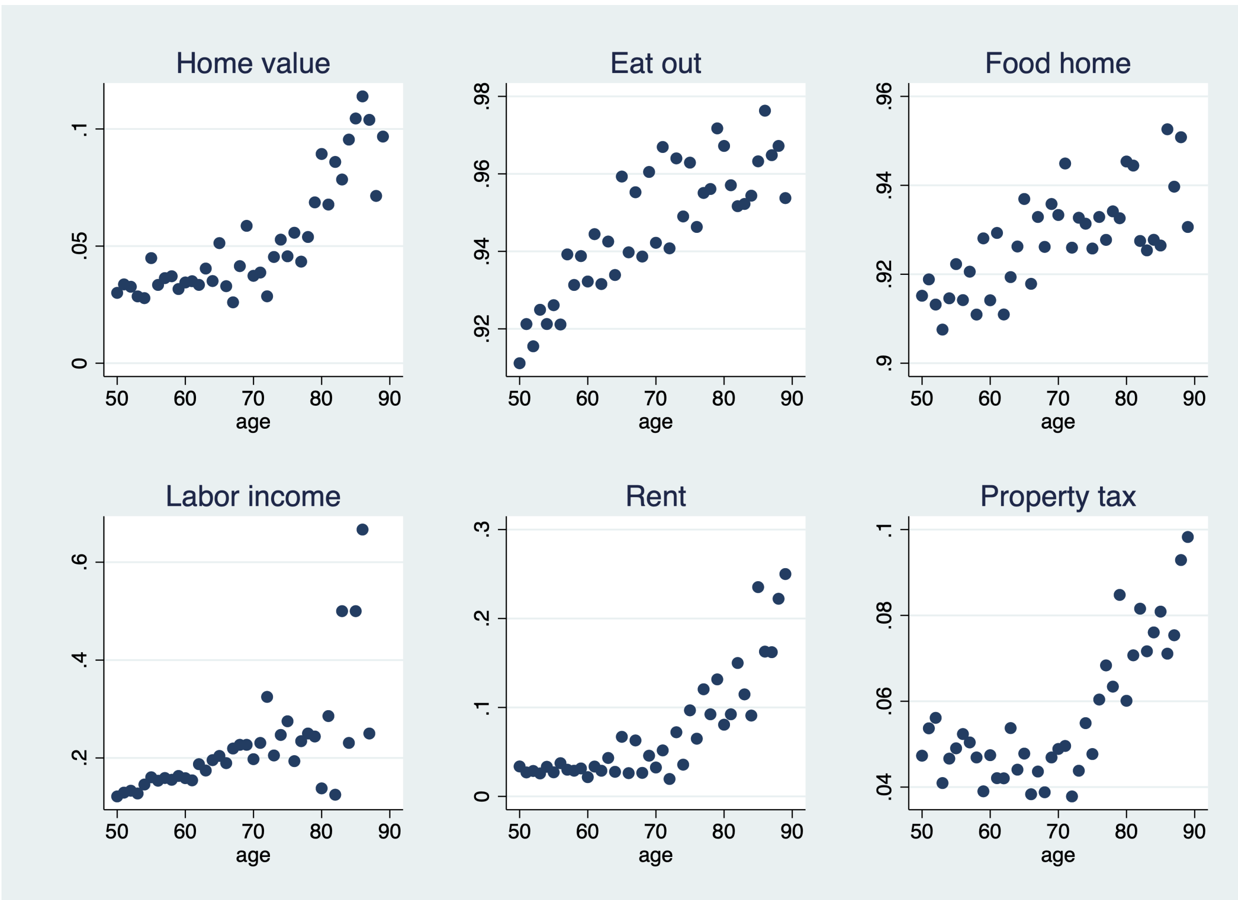
All the panels restrict the household head respondents aged between 50 and 89 who participated in the PSID survey from 1999 to 2019. Each panel summarizes the level of round on the numerical response against age in years. Note that Gideon, Helppie-McFall, and Hsu (2017) define the level of rounding as $\left(\frac{\text{the number of total digits} - \text{the number of significant digits}}{\text{the number of total digits} - 1}\right)$. The higher the level of round, the more the respondents round up the responses. The panel title indicates the question.

Figure A.3: Maximal rounding by age in PSID



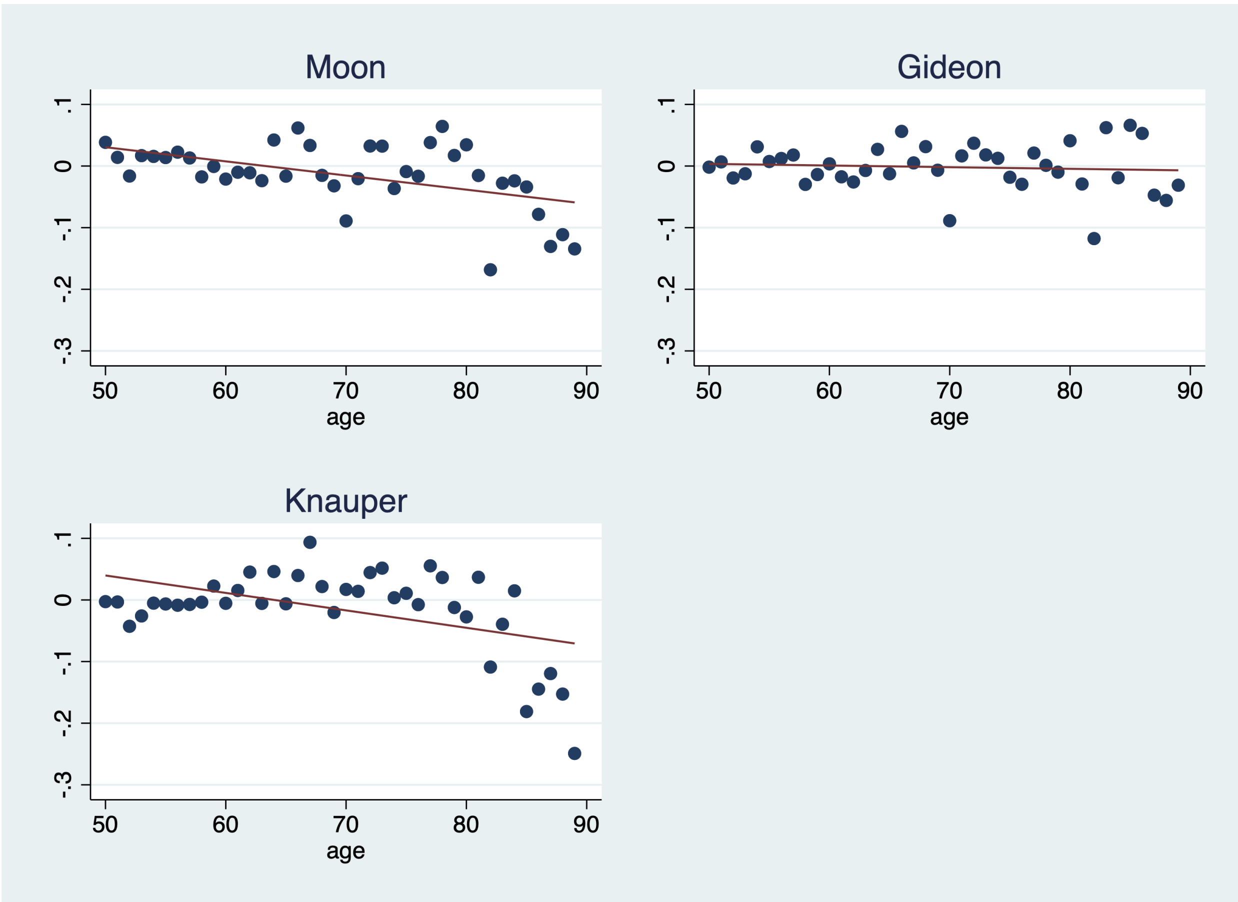
All the panels restrict the household head respondents aged between 50 and 89 who participated in the PSID survey from 1999 to 2019. Each panel summarizes whether one uses the maximal rounding conditioning that the respondent provides a numerical response. The maximal rounding indicates the numerical response format in which the number is rounded to the precision of 1. That is, the leftmost digit is any number, and the rest of the digit positions are zeros, for example, 2,000 or 100,000. Each panel represents the share of the maximal rounding out of numerical responses by age. The panel title indicates the question.

Figure A.4: Opt-out trend by age in PSID



All the panels restrict the household head respondents aged between 50 and 89 who participated in the PSID survey from 1999 to 2019. Each panel summarizes whether the respondents choose the opt-out responses. The HRS provides the opt-out options for the open-ended numerical questions, such as for *Don't know* and *Refused to answer*. Each panel represents the share of opt-out responses by age. The panel title indicates the question.

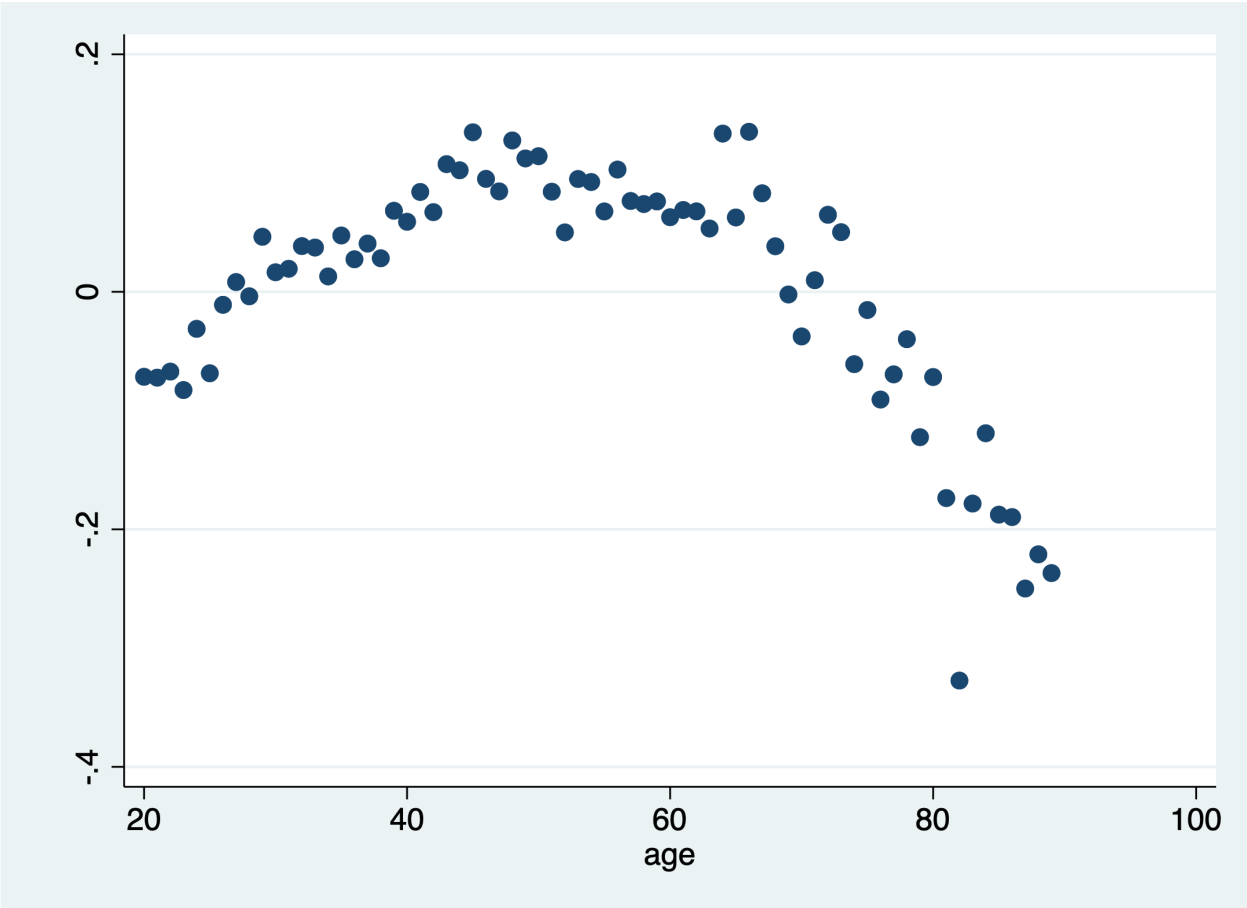
Figure A.5: Age trend of the cognitive proxies controlling for individual effects in the PSID



Each proxy is residualized by the non-time varying controls within the same respondents. Then, the corresponding residuals are plotted against age in years. The sample is limited to the household heads aged between 50 and 89 who participated in the survey from 1999 to 2019. Each panel presents the average of the residual proxies against age in years. All the proxies are standardized.

I characterize each open-ended response in the following way.
Moon: assign 0 to opt-out response; 1 to the maximal rounding; 2 to other numerical response.
Gideon: calculate the level of round as $\left(\frac{\text{the number of total digits}-\text{the number of significant digits}}{\text{the number of total digits}-1}\right)$. I use complement to 1 of the level of round, because I want the higher value to imply higher cognitive ability.
Knäuper: assign 0 to opt-out response; 1 to other numerical response.
After characterizing every response, I take the average of them.
*opt-out responses indicate either *Don't know* or *Refused to answer*. For more details, see Table 6.

Figure A.6: Age trend of the Moon proxy



The figure plots the average the cognitive proxy *Moon* against age in years, constructed from the PSID data. The figure restricts the household heads respondents aged between 20 and 89 who participated in the survey from 1999 to 2019. The proxy is standardized.